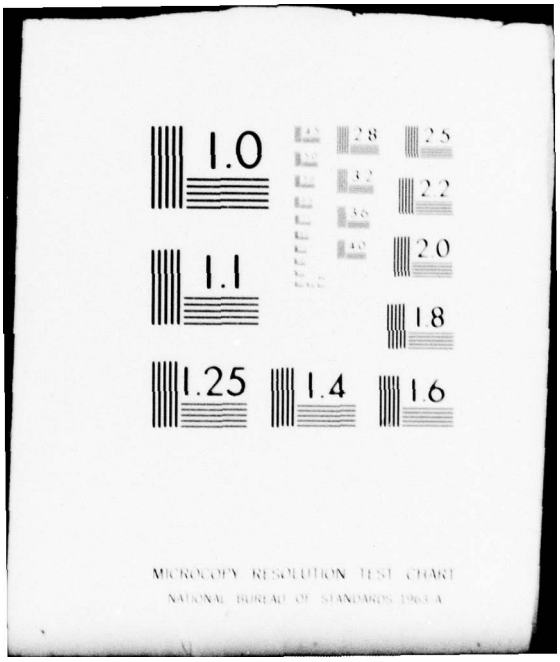
AD-A068 660 STANFORD UNIV CALIF INFORMATION SYSTEMS LAB F/G 17/4
RESEARCH ON ADAPTIVE ANTENNA TECHNIQUES III. PART 1. THE SOFT C--ETC(U)
JAN 79 B WIDROW, R CHESTEK, T SAXE N00019-78-C-0276
UNCLASSIFIED NL

| OF |
AD A068660

END
DATE FILMED
6-79
DDC


1.0
1.1
1.25
1.4
1.6
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.8
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

INFORMATION SYSTEMS LABORATORY

STANFORD ELECTRONICS LABORATORIES
DEPARTMENT OF ELECTRICAL ENGINEERING
STANFORD UNIVERSITY · STANFORD, CA 94305

12

RESEARCH ON ADAPTIVE ANTENNA TECHNIQUES III

LEVEL II

FINAL REPORT

by

B. Widrow

Richard Chestek

Timothy Saxe

January 31, 1979

This research was supported by the
Naval Air Systems Command of the
Department of Defense under
Contract N00019-78-C-0276

APPROVED FOR PUBLIC RELEASE:
DISTRIBUTION UNLIMITED

Principal Investigator: Bernard Widrow (415)497-4949

Name of Contractor: Stanford University

Effective Date of Contract: February 1, 1978 to January 31, 1979

Short Title of Work: Research on Adaptive Antenna Techniques III

Amount of Contract: $59,989

FINAL REPORT

PART 1:

THE SOFT CONSTRAINED LMS ALGORITHM

PART 2:

AN IMPROVED ADAPTIVE SEPARATOR

Final Rept.

by

B. Widrow
Richard Chestek
Timothy Saxe

31 Jan 79

89 p.

This research was supported by the Naval Air Systems Command of the Department of Defense under Contract N00019-78-C-0276

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Naval Air Systems Command or the U. S. Government.

APPROVED FOR PUBLIC RELEASE:
DISTRIBUTION UNLIMITED

AD A067209

FOREWORD

This research is concerned with separation of desired signals and jamming signals based on their relative power levels. The jammers are assumed to be strong compared to the desired signals, otherwise they are of no concern since other methods can be used to eliminate them. In PART 1, we develop "soft constraint" adaptive algorithms that are suitable for power separation with adaptive arrays. Signals and jammers of various power levels can be separated as long as they differ in angle of arrival and/or frequency band. In PART 2, we develop single channel power separators with sharper thresholds than have been previously possible. Such systems can separate signals and jammers of various power levels as long as they differ in frequency band. The goal is to study sharp threshold phenomena in the simpler context of single channel operation with the expectation of final application to multi-channel adaptive arrays.

PART 1. THE SOFT CONSTRAINED LMS ALGORITHM

1.1. Introduction

The Widrow Hoff Least Mean Square (LMS) algorithm [1,2] is a well known method of adapting the weights of a linear combiner (Fig. 1.1) to perform least mean square estimation. The adaptation allows the filter to approach the optimum set of parameters based on the current statistics of the inputs to the linear combiner. A major application has been in the field of adaptive filters, where the inputs to the linear combiner are samples of time data typically the outputs of a tapped delay line (Fig. 1.2).

Adaptive filters using the LMS algorithm have been proposed for use in many applications [3-7]. However in some situations it has been necessary or desirable to modify the algorithm [8-12]. Frost [9] proposed the addition of hard constraints. The hard constraint forces the set of weights to obey a set of linear equalities. This modification of the LMS algorithm has found application in the field of adaptive antenna arrays, to force the array's gain to be exactly unity to a signal arriving from a specified direction, but allowing signals arriving from other directions to be greatly attenuated.

Another modification of the LMS algorithm is the "leaky" LMS algorithm. This algorithm has a leak factor built in such that in the absence of any inputs the weights will decay to zero. This form has been proposed independently by several researchers [11-14]. Using the property that the leak is equivalent to introducing a white noise to the input of the filter, Treichler [11] proposed using this equivalent

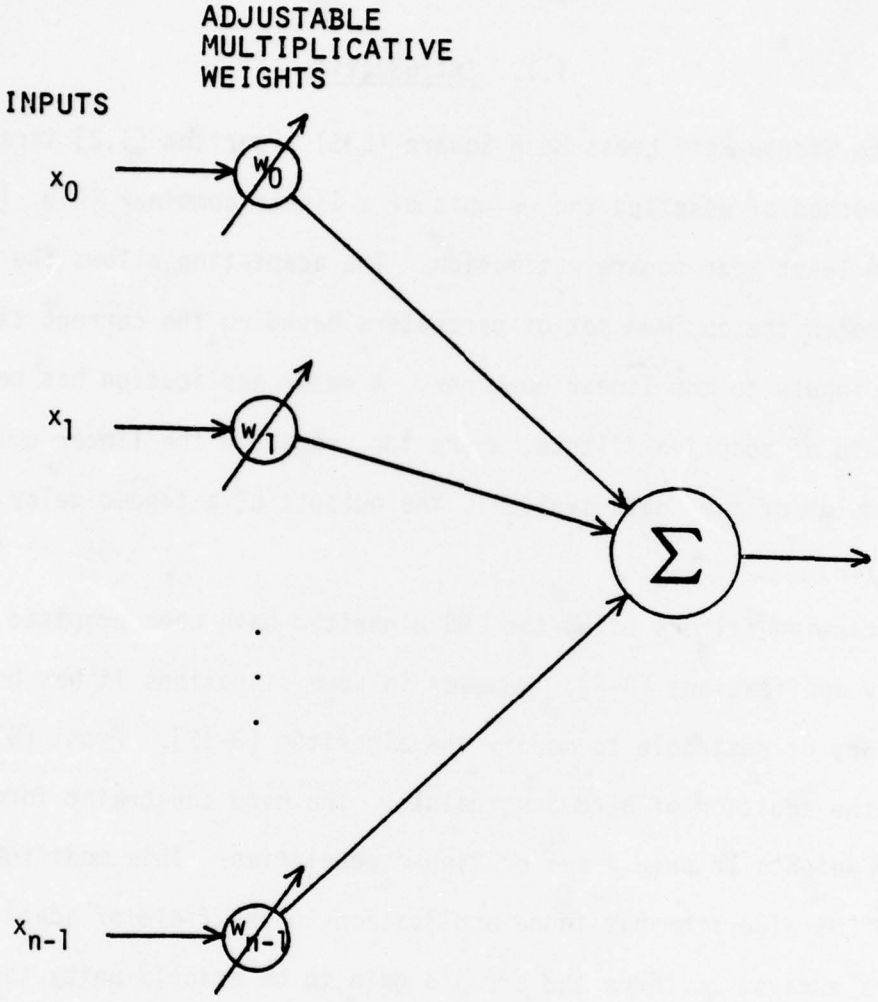


FIG. 1.1 LINEAR COMBINER

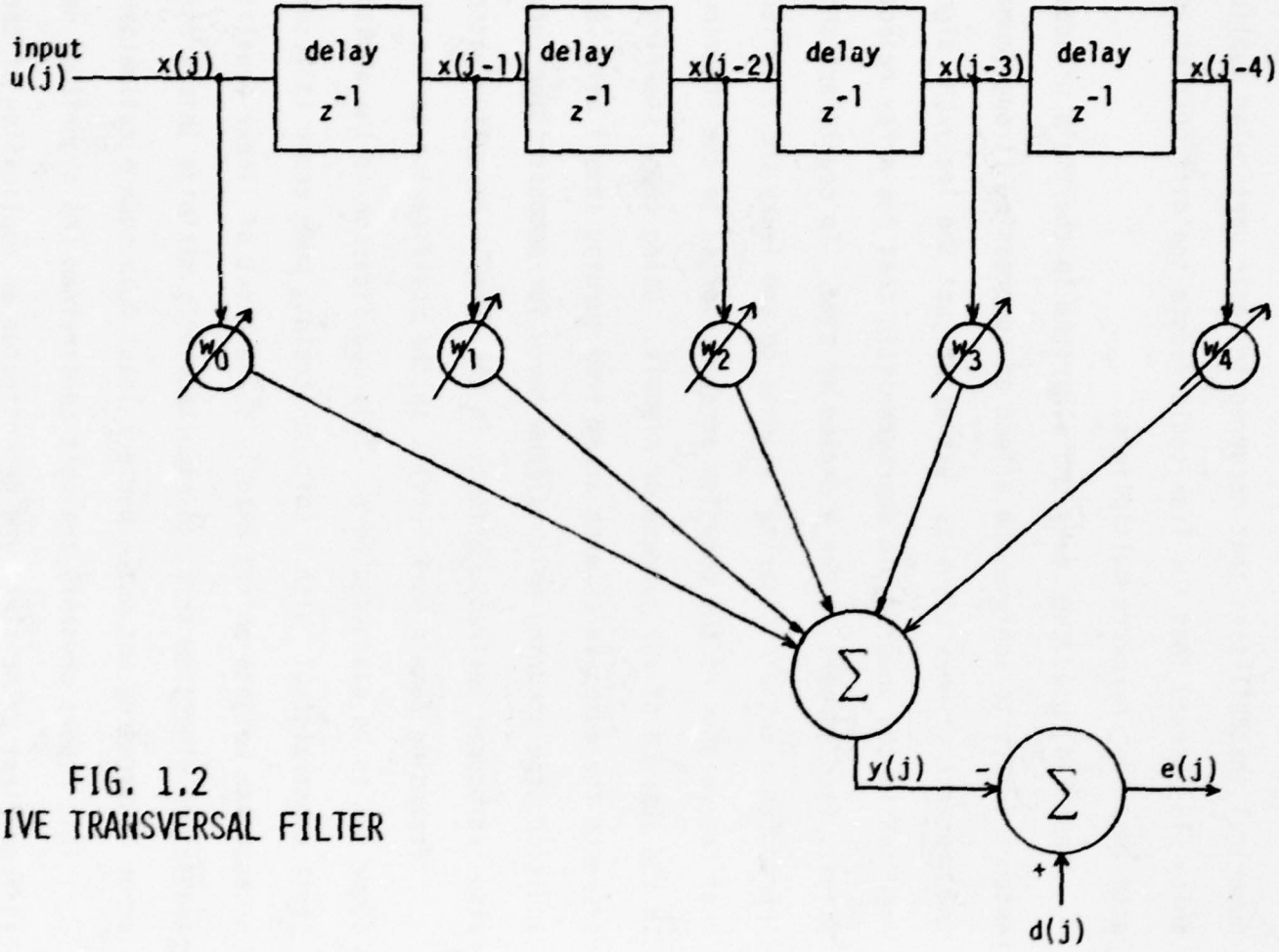


FIG. 1.2
ADAPTIVE TRANSVERSAL FILTER

white noise to modify the characteristics of an adaptive line enhancer in a desired fashion. Ahmed et al [12] used the effect to reduce numerical instabilities that occurred in their particular application. White [13] showed that the leak could reduce the difficulties associated with imperfect hardware multipliers.

Zahm [14] used the leaky LMS algorithm in the field of adaptive antenna arrays to achieve the effect of suppressing strong jammers in the presence of weak signals. But using just the leaky LMS algorithm resulted in the undesirable characteristic that the array rejected all signals (and jammers) after a period of time. To counteract this, Zahm introduced a set of "steering" weights to the leaky LMS algorithm so that the weights of the adaptive array converged to the steering weights in the absence of any jammers or signals. Using these steering weights prevents the adaptive antenna array from turning itself off. But in addition, the steering weights Zahm chose for demonstration purposes also introduced desirable effects in the array's reception pattern.

Extending Zahm's work results in the modification to the LMS algorithm to be discussed here. This modification is the addition of "soft constraints." With a soft constraint, some error is incurred because the weights do not exactly solve a set of linear equalities. The adaptation algorithm tries to simultaneously minimize this error and the error incurred by not doing perfect least mean square estimation.

This report presents the soft constrained LMS algorithm, deriving it from first principles, and demonstrates an application of the soft constraint algorithm in the area of adaptive antenna arrays.

1.2 Development of the Soft Constrained LMS Algorithm

This section develops the soft constrained LMS algorithm discussed in the introduction. The development begins from first principles, by constructing a performance criterion. Once the performance criterion is obtained, the optimum weight vector which minimizes this performance criterion is derived. Finally, the soft constrained LMS algorithm is developed as a steepest descent search of the performance surface for its minimum point, which occurs at the optimum weight vector.

1.2.1 Definitions and Terminology

Although applicable to any linear combiner, this work will assume for ease of discussion that the soft-constrained LMS algorithm will be used for an adaptive filter. Figure 1.2 portrays an adaptive filter. Definitions and terminology related to the adaptive filter are now presented.

A sampled time sequence u(j) is the input to an n-1 tap transversal filter. The n weights w_i (i=1, ..., n) are changed by the adaptation algorithm as time progresses. The output y(j) of the filter is compared against a time sequence d(j), which will be called the desired signal. The purpose of the filter is to provide an output y(j) which is an estimate of the desired signal d(j). Thus y(j) will be called an estimate of d(j). The difference between desired signal and the estimate will be called the error signal, e(j).

The time sequence u(j) may be the sum of three types of signals. The signal may consist of noise, or may be a deliberately produced sequence of no use in forming the estimate (interferers), or may be a sequence helpful in estimating d(j).

The values at the taps of the transversal filter at time j will be denoted by the vector:

$$X(j) = [u(j)\ u(j-1)\ \ldots\ u(j-n+1)]^T \qquad (1\text{-}1)$$

X(j) will be called the data vector.

The set of weights will also be written in vector form as:

$$W = [w_1\ w_2\ \ldots\ w_n]^T \quad . \qquad (1\text{-}2)$$

Thus, the filter output y(j) is expressed in vector form as

$$y(j) = X(j)^T W = W^T X(j) \quad . \qquad (1\text{-}3)$$

The error signal is simply

$$e(j) = d(j) - y(j) \quad . \qquad (1\text{-}4)$$

1.2.2 The Performance Criterion

One of the goals of the adaptive filter is to form a good estimate of the desired signal. A standard measure of the "goodness" of an estimate is the mean square error [15-17]. This is written as $E\{e^2(j)\}$ where $E\{a\}$ denotes the expected value of a. A good estimate will be one which minimizes the performance criterion $E\{e^2(j)\}$. The minimization is performed by selection of a suitable weight vector, W.

Expanding the mean square error in terms of W yields:

$$\begin{aligned} E\{e^2(j)\} &= E\{(d(j) - y(j))^2\} \\ &= E\{d(j)^2 - 2d(j)y(j) + y(j)^2\} \\ &= E\{d(j)^2 - 2d(j)\ X(j)^T W + W^T X(j) X(j)^T W\} \quad . \end{aligned} \qquad (1\text{-}5)$$

Denote the cross correlation between the data vector X(j) and the desired signal d(j) by P(j):

$$P(j) = E\{d(j)X(j)\} \quad . \tag{1-6}$$

Also denote the autocorrelation matrix of the input data by $\underline{R}(j)$:

$$\underline{R}(j) = E\{X(j)X(j)^T\} \quad . \tag{1-7}$$

Using these definitions in Eq. 1-5 gives an expression for the mean square error as a function of the weight vector W:

$$E\{e^2(j)\} = E\{d(j)^2\} - 2W^TP(j) + W^T\underline{R}(j)W \quad . \tag{1-8}$$

Clearly, the mean square error is a quadratic function of the weights. Thus it may be visualized as a parabolic surface in a hyperspace of dimension n. $\underline{R}(j)$ is positive semidefinite, since it is an autocorrelation matrix. If $\underline{R}(j)$ is nonsingular (and therefore positive definite), then there is a unique weight vector which minimizes the mean square error. However, if $\underline{R}(j)$ is singular, then there is a set of weight vectors, all of which yield the same minimum mean square error.

But in addition to minimizing the mean square error, it is also desirable to keep close to a set of linear constraints on the weight vector. If possible, a set of m linear constraints should be satisfied:

$$A_zW = h_z \qquad (z=1,\ldots,m) \tag{1-9}$$

where A_z is a row vector, and h_z is the value the linear combination of the weights should take on. However, it may be that it is impossible for all of the constraints to be satisfied simultaneously, or satisfaction of a constraint may result in a value of mean squared error unacceptably high to the designer. Thus, the equality of Eq. 1-9 may not always be true at the most desirable weight vector. Define the error created by failure to satisfy the z^{th} constraint by:

$$e_{cz} = A_z W - h_z \tag{1-10}$$

which is the difference between the actual value of the linear combination of the weights, and the desired value. e_{cz} is called the constraint error. To measure the performance in meeting the constraints, form a weighted sum of the squares of the constraint errors:

$$\sum_{z=1}^{m} b_z (e_{cz})^2 = \sum_{z=1}^{m} b_z (A_z W - h_z)^2 \tag{1-11}$$

where the b_z are a set of positive constants chosen by the designer. The greater the value of an element b_z, the more the error e_{cz} will affect the total error. Thus the designer specifies the relative importance of each constraint by selection of b_z. This controls the "stiffness" of the constraint. Examples of the effect of varying b_z are shown in a later section. Equation 1-11 is written in matrix form as:

$$(\underline{A}W - H)^T \underline{B} (\underline{A}W - H) \tag{1-12}$$

where $\underline{A}$ is the matrix composed of the vectors A_z:

$$\underline{A} = [A_1^T \; A_2^T \; \ldots \; A_m^T]^T \quad . \tag{1-13}$$

B is the diagonal matrix with diagonal elements b_z:

$$\underline{B} = \text{diag}[b_1, b_2, \ldots, b_m] \tag{1-14}$$

and H is a vector composed of the individual desired constraint values h_z:

$$H = [h_1 \; h_2 \; \ldots \; h_m]^T \quad . \tag{1-15}$$

Thus, $\underline{A}$ is an mxn matrix, $\underline{B}$ is an mxm matrix, and H is a vector of dimension m. Note that the number of constraints does not have to equal

the number of weights. If n>m, there are more weights than constraints, which is called an underconstrained problem. Alternately, if n<m, there are more constraints than weights, which is an overconstrained problem. When n=m, the number of constraints and the number of weights are equal; this is called an exactly constrained problem.

Note from expression 1-12 that this summed squared constraint error is also a quadratic function of the weight vector, like the mean squared error. Because the summed squared criterion error is the sum of squared quantities weighted by nonnegative factors, it is bounded from below by zero, and therefore cannot decrease without bound. Therefore, a minimum value exists. This minimum summed squared criterion error will be obtained by a unique weight vector in some cases, and obtainable by a whole set of weight vectors in other cases. The condition for the minimum to be obtained by a unique weight vector is that the matrix $\underline{A}$ be of rank n (it must have n linearly independent rows).

Because of the possibility of not satisfying all or any of the constraints in a given problem, these constraints are called soft constraints, as opposed to hard constraints, which must be satisfied absolutely.

Now performance criteria have been defined for both the signal estimation problem (the mean square error) and the constraint problem (sum of squared constraint errors). The performance criterion for a soft constrained signal estimation problem is defined as the sum of these two individual performance criteria:

$$p(j) = E\{e^2(j)\} + (\underline{A}W - H)^T \underline{B}(\underline{A}W - H) \quad . \qquad (1\text{-}16)$$

The goal of the soft constrained LMS algorithm is to find the weight vector W which minimizes the performance criterion p(j).

Since p(j) is the sum of two expressions which are quadratic in the weight vector W, p(j) itself is also a quadratic function of the weight vector.

1.2.3 Optimum Solution

In this sub-section an expression is found for the optimum weight vector, which is defined as the unique weight vector which minimizes the performance criterion p(j). The condition under which the minimum p(j) occurs with a non-unique weight vector is also determined.

Any weight vector W which minimizes the performance criterion p(j) causes the gradient of p(j) to become zero. An expression for the gradient of p(j) with respect to W is:

$$\nabla_W p(j) = \nabla_W E\{e^2(j)\} + \nabla_W[(\underline{A}W - H)^T \underline{B}(\underline{A}W - H)] \qquad (1\text{-}17)$$

Analyzing the first term yields:

$$\begin{aligned}
\nabla_W E\{e^2(j)\} &= \nabla_W E\{(d(j)-y(j))^2\} \\
&= \nabla_W E\{(d(j)-W^T X(j))^2\} \\
&= \nabla_W E\{d(j)^2-2d(j)W^T X(j)+W^T X(j)X(j)^T W\} \\
&= -2(E\{d(j)X(j)\}-E\{X(j)X(j)^T\}W)^T \\
&= -2(P(j)-\underline{R}(j)W)^T \qquad (1\text{-}18)
\end{aligned}$$

This first term of the gradient is attributable to the mean square error term of the performance criterion. It is the gradient used in developing the LMS algorithm.

Analyzing the second term of Eq. 1-17 yields:

$$\nabla_W(\underline{A}W - H)^T\underline{B}(\underline{A}W - H) = [2\underline{A}^T\underline{B}(\underline{A}W - H)]^T \quad . \qquad (1\text{-}19)$$

This second term is due entirely to the soft constraints imposed on the problem by the filter designer.

The gradient is the sum of the two terms:

$$\nabla_W^T p(j) = -2(P(j)-\underline{R}(j)W)+2\underline{A}^T\underline{B}(\underline{A}W - H) \quad . \qquad (1\text{-}20)$$

The optimum value for W is found by setting the gradient equal to zero and solving for W:

$$(\underline{R}(j)+\underline{A}^T\underline{B}\underline{A})W = P(j)+\underline{A}^T\underline{B}H \quad . \qquad (1\text{-}21)$$

It is seen that the necessary condition for the optimum (minimum) performance to occur at a unique weight vector is that the matrix $\underline{R}(j)+\underline{A}^T\underline{B}\underline{A}$ be nonsingular. Under this condition, the unique optimum weight vector (denoted W_{opt}) is:

$$W_{opt} = (\underline{R}(j)+\underline{A}^T\underline{B}\underline{A})^{-1}(P(j)+\underline{A}^T\underline{B}H) \quad . \qquad (1\text{-}22)$$

Note that it is not necessary for either $\underline{R}(j)$ or $\underline{A}^T\underline{B}\underline{A}$ to be nonsingular by themselves. In fact, one of the uses of the soft constrained LMS algorithm is in situations where the data autocorrelation matrix $\underline{R}(j)$ is singular (or possibly just ill-conditioned), and a set of soft constraints is therefore generated to yield a unique optimum weight vector [12].

1.2.4 An Assumption of Stationarity

The remainder of this work will assume that the signals d(j), and u(j) are both stationary. In this case, the statistics P(j), $\underline{R}(j)$, and

the performance criterion p(j) are constant, so the time index j will be dropped, denoting them now by P, $\underline{R}$, and p.

Although not noted above, since p was previously a function of time, the optimum weight vector W_{opt} was also a function of time. Now, due to the assumption of stationarity, W_{opt} is independent of time, and has the value:

$$W_{opt} = (\underline{R}+\underline{A}^T\underline{B}\underline{A})^{-1} (P+\underline{A}^T\underline{B}H) \quad . \qquad (1\text{-}23)$$

1.2.5 Determination of the Optimum Weight Vector by Gradient Search

Calculation of the optimum weight vector using Eq. 1-23 is not always feasible, even when all of the quantities are known. This may be due to the size of the filter, or numerical difficulties due to properties of the matrices. Thus alternative approaches to calculation have come into existence. A common technique is to make successive approximations to the optimum weight vector. Given one estimate of the weight vector, denoted by W(j), the next estimate, W(j+1), is generated from W(j) governed by how well W(j) satisfies Eq. 1-21.

The particular technique of successive approximation used in this research is called gradient search [18]. Very simply, the gradient of the performance surface is calculated for the current setting of the weight vector W(j). Since the gradient specifies the direction of weight vector change which will increase the performance function most rapidly, and the goal is to reduce the performance function, the next estimate of the optimum weight vector is obtained by moving from the current estimate in the direction opposite to that of the gradient, and a distance proportional to the magnitude of the gradient:

$$W(j+1) = W(j) - \mu\nabla_W p \qquad (1\text{-}24)$$

μ is a positive constant chosen by the filter designer. Selection of μ is subject to considerations discussed in the next section.

Using Eq. 1-20 for $\nabla_W p$ (and dropping the subscript j on $\underline{R}(j)$ and P(j) due to the assumption of stationarity) gives the update equation:

$$W(j+1) = W(j) + 2\mu(P-\underline{R}W(j)) - 2\mu\underline{A}^T\underline{B}(\underline{A}W(j)-H) \quad . \qquad (1\text{-}25)$$

Repeated use of this update equation will cause the estimate of the optimum weight vector to approach W_{opt}, the actual optimum, provided μ is small enough.

1.2.6 The Soft Constrained LMS Algorithm

The algorithm described above for finding the optimum weight vector W_{opt} is applicable only if all quantities are known. This may not always be true. In particular, if the statistics of the input signal u(j) are not perfectly known, then the quantities P and $\underline{R}$ are unknown. This can occur when a known signal is subject to additive noise, or is passed through a filter whose characteristics are not perfectly known, or subject to distortion. However, it is still desirable to perform signal estimation subject to soft constraints in this case. Accordingly, the update equation (Eq. 1-24) is modified to handle the situation. In particular, P and $\underline{R}$ will be replaced by appropriate estimates. The estimates chosen must be dependent upon u(j), so that the estimates are based on true statistics, and not an a-priori guess. The estimates chosen are:

$$\hat{P} = d(j)X(j)$$

$$\hat{\underline{R}} = X(j)X(j)^T \quad . \qquad (1\text{-}26)$$

It is easily shown that these estimates are unbiased:

$$E\{\hat{P}\} = E\{d(j)X(j)\} = P$$

$$E\{\hat{\underline{R}}\} = E\{X(j)X(j)^T\} = \underline{R} \quad . \tag{1-27}$$

Thus the update algorithm becomes

$$W(j+1) = W(j) - \mu\hat{\nabla}_W p \tag{1-28}$$

where an estimate of the gradient is now used in place of the true gradient:

$$\hat{\nabla}_W p = -2(\hat{P}-\hat{\underline{R}}W(j)) + 2\underline{A}^T\underline{B}(\underline{A}W(j)-H) \quad . \tag{1-29}$$

Substituting the expressions 1-26 into the expression for the gradient estimate above results in:

$$\hat{\nabla}_W p = -2(d(j)X(j)-X(j)X(j)^T W(j)) + 2\underline{A}^T\underline{B}(\underline{A}W(j)-H) \quad . \tag{1-30}$$

Noting that $X(j)^T W(j)$ is $y(j)$ and substituting:

$$\begin{aligned}\hat{\nabla}_W p &= -2(d(j)X(j)-X(j)y(j)) + 2\underline{A}^T\underline{B}(\underline{A}W(j)-H) \\ &= -2(d(j)-y(j))X(j) + 2\underline{A}^T\underline{B}(\underline{A}W(j)-H) \quad .\end{aligned} \tag{1-31}$$

Finally since $d(j)-y(j)$ is the error $e(j)$:

$$\hat{\nabla}_W p = -2e(j)X(j) + 2\underline{A}^T\underline{B}(\underline{A}W(j)-H) \quad . \tag{1-32}$$

Writing the update equation as one expression results in:

$$W(j+1) = W(j) + 2\mu e(j)X(j) - 2\mu\underline{A}^T\underline{B}(\underline{A}W(j)-H) \tag{1-33}$$

or

$$W(j+1) = (\underline{I}-2\mu\underline{A}^T\underline{B}\underline{A})W(j) + 2\mu e(j)X(j) + 2\mu\underline{A}^T\underline{B}H \quad . \tag{1-34}$$

This is defined here as the soft constrained LMS algorithm.

It is interesting to examine the gradient estimate a little more closely. By comparing the gradient estimate of Eq. 1-29 with the true gradient in Eq. 1-20 it can be seen that only the gradient term due to mean square error is approximated. The gradient term due to the soft constraints is still calculated perfectly, entirely from knowledge of the soft constraints and the current weight vector.

1.3 An Application to Adaptive Antenna Arrays

This section demonstrates an application of the soft constrained LMS algorithm to adaptive antenna arrays. The soft constraints are used to affect the shape of the antenna array's directivity pattern.

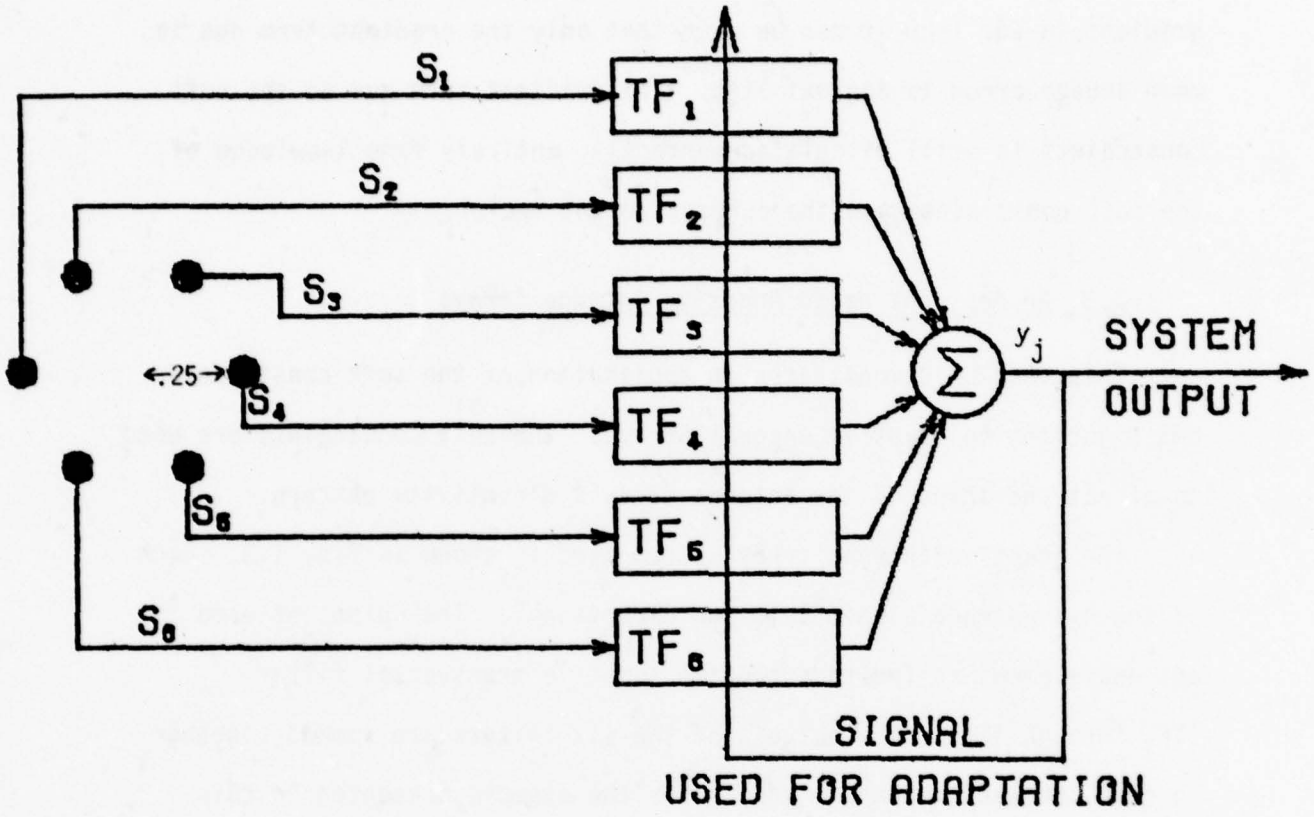
The adaptive antenna array system used is shown in Fig. 1.3. Each of the six antenna elements is omnidirectional. The output of each antenna element is fed to a two tap adaptive transversal filter (TF_1 through TF_6). The outputs of the six filters are summed together to form the system output, y(j). For the example presented in this section, it is assumed that no desired signal d(j) is available.

The weight vector of the antenna array system is the vector constructed by stacking the weight vectors of the individual adaptive filters together. Denoting the weight vector of transversal filter k at time j by $W_k(j)$, the weight vector W(j) of the entire system is:

$$W(j) = [W_1^T(j)\ W_2^T(j)\ \ldots\ W_6^T(j)]^T \quad . \tag{1-35}$$

The data vector X(j) for the entire system is constructed similarly.

The soft constraints are used to specify desirable antenna gains in a particular direction at a specified frequency.



SPEED OF PROPAGATION = 1

SAMPLING INTERVAL = .125

FIG. 1.3 AN ADAPTIVE ANTENNA ARRAY

Imagine the antenna array receiving a sinusoid of power C_z^2 at frequency ω_z from a specified direction θ_z. Denote the signals at the input to transversal filter TF_k by the phasor notation $C_z\exp\{i(\omega_z jT+\phi_{zk})\}$ where ϕ_{zk} is the phase difference of the signal between sensor k and some reference point. ϕ_{zk} is a function of both the angle of arrival of the signal (θ_z) and the antenna geometry. In this case the data vector is:

$$X(j) = \left| \begin{array}{l} C_z\exp\{i(\omega_z jT+\phi_{z1})\} \\ C_z\exp\{i(\omega_z (j-1)T+\phi_{z1})\} \\ \hline \vdots \\ \hline C_z\exp\{i(\omega_z jT+\phi_{z6})\} \\ C_z\exp\{i(\omega_z (j-1)T+\phi_{z6})\} \end{array} \right| \begin{array}{l} \text{(data in } TF_1\text{)} \\ \\ \\ \\ \\ \text{(data in } TF_6\text{)} \end{array} \qquad (1\text{-}36)$$

Then the array output is $X^T(j)W(j)$. The array gain to this signal is $X^T(j)W(j)/C_z\exp\{i\omega_z jT\} = (X^T(j)/C_z\exp\{i\omega_z jT\})W(j)$. Define the row vector A_z by $X^T(j)/C_z\exp\{j\omega_i jT\}$. Then

$$A_z^T = \left| \begin{array}{l} \exp\{i\phi_{z1}\} \\ \exp\{i(\phi_{z1}-\omega_z T)\} \\ \hline \vdots \\ \hline \exp\{i\phi_{z6}\} \\ \exp\{i(\phi_{z6}-\omega_z T)\} \end{array} \right| . \qquad (1\text{-}37)$$

The array gain to signal z at time j is thus $A_z W(j)$. Suppose it is desirable to have the array gain to this signal be $D_z \exp\{i\eta_z\}$. Then the constraint can be written as

$$A_z W(j) = D_z \exp\{i\eta_z\} \tag{1-38}$$

and can be made a soft constraint. However, note that W(j) is a set of real weights, while A_z and $D_z \exp\{i\eta_z\}$ are complex. This constraint can still be realized by separating it into the real and imaginary parts:

$$\mathrm{Re}\{A_z\} W(j) = \mathrm{Re}\{D_z \exp\{i\eta_z\}\} \tag{1-39}$$

$$\mathrm{Im}\{A_z\} W(j) = \mathrm{Im}\{D_z \exp\{i\eta_z\}\} \quad . \tag{1-40}$$

This process yields two constraints which can be used as soft constraints. Thus the antenna array will attempt to keep a complex gain of $D_z \exp\{i\eta_z\}$ in direction θ_z at freqeuncy w_z. These two constraints will be used for soft constraints, and thus the array gain can vary from the specified gain ($D_z \exp\{i\eta_z\}$).

This procedure can be followed for several different sinusoids, at the same or different frequencies, yielding a set of constraints. Let the set of constraint vectors (A_z) be formed into a matrix $\underline{A}$, and let the gain specifications be stacked into a corresponding vector H. Then the set of soft constraints is

$$\underline{A}W = H \quad . \tag{1-41}$$

The soft constraints will be weighted by constants b_z, which are used to compose the diagonal matrix $\underline{B}$.

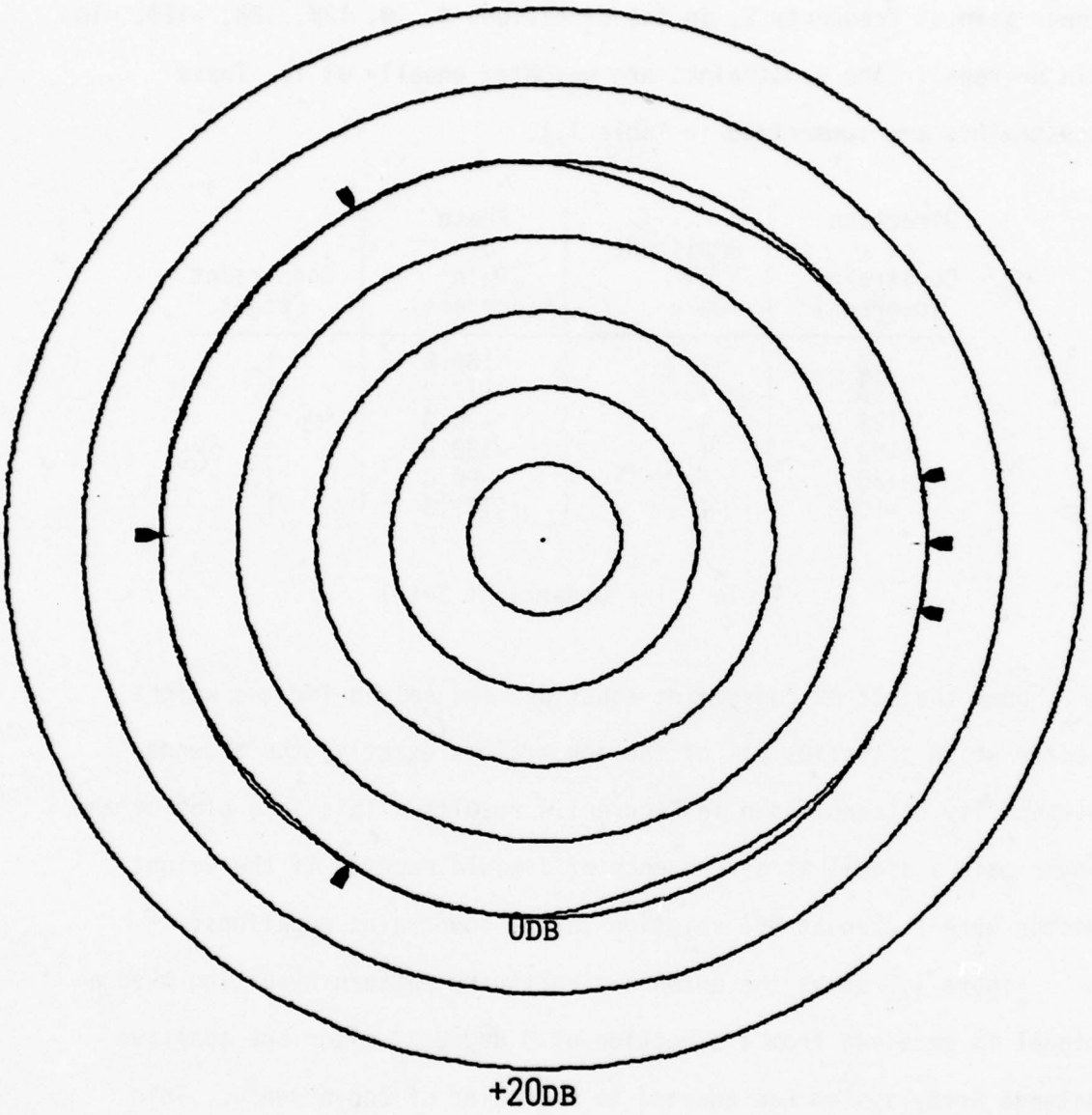
The first example of the use of soft constraints is shown in Figures 1.4 through 1.8. The soft constraints are specified to have unity power gain at frequency 2, in the directions Ø, 1Ø, 12Ø, 18Ø, -12Ø, -1Ø (in degrees). The constraints are weighted equally at 1. These constraints are summarized in Table 1.1.

Direction of Constraint (degrees)	Amplitude of Gain	Phase of Gain (degrees)	Constraint Weight
Ø.	1.	18Ø.Ø	1.
1Ø	1.	177.3	1.
12Ø	1.	-9Ø.Ø	1.
18Ø	1.	-18Ø.Ø	1.
-12Ø	1.	-9Ø.Ø	1.
-1Ø	1.	177.3	1.

Table 1.1 - Constraint Set 1

When the set of constraint equations are solved for the weight vector which satisfies all of the constraints exactly, the antenna directivity pattern shown in Figure 1.4 results. This is a plot of the power gain a signal at a frequency of 2 would receive if the weight vector were frozen at the solution to the constraint equations.

Figure 1.5 shows the antenna directivity pattern resulting when a signal is received from a direction of Ø degrees, after the adaptive antenna array system has adapted to the point of convergence. This example (and all others in this section) also has an isotropic noise field impinging on the antenna array. The noise power at each antenna element is Ø.1. In this case, since there is no desired signal, the mean squared error is just the system output power, so the goal of the soft constrained LMS algorithm is to minimize the system output power



◄ - CONSTRAINT (ALL CONSTRAINTS EQUALLY WEIGHTED)

FIG. 1.4 ANTENNA ARRAY DIRECTIVITY PATTERN
DETERMINED BY SOFT CONSTRAINTS

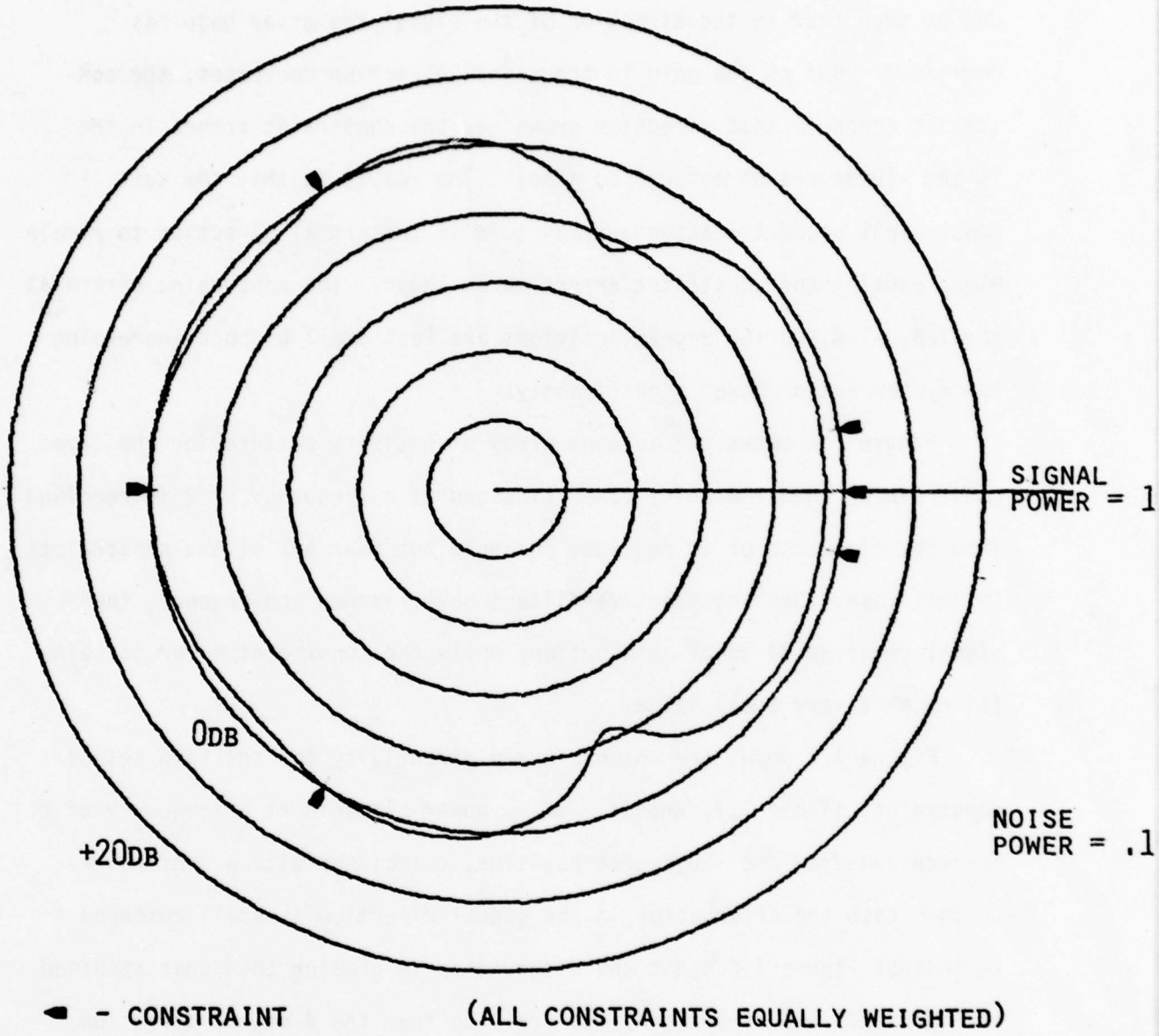


FIG. 1.5 ANTENNA ARRAY DIRECTIVITY PATTERN WITH SOFT CONSTRAINTS AND ONE SIGNAL

while trying to keep the array gain close to the constraint values. It can be seen that in the direction of the signal the array gain has decreased. But as the gain in the signal direction decreases, the constraint error in that direction grows (as the constraint errors in the 1Ø and -1Ø degree directions do also). The result is that the soft constraints cause the antenna array gain in the signal direction to remain high, causing the constraint errors to be lower. The constraint errors at the 12Ø, -12Ø and 18Ø degree positions are kept small without increasing the system output power significantly.

Figure 1.6 shows the antenna array directivity pattern for the same constraints, when the unity power sinusoid at a frequency of 2 is received from the direction of 6Ø degrees, an angle not near any of the constraints. In this case, when the adaptive filters have reached convergence, the signal receives 3Ø dB of attenuation, while the constraint error is maintained at a very small value.

Figure 1.7 shows the antenna array directivity for the same set of constraints (Table 1.1) when the unity power sinusoid at a frequency of 2 is received from the 12Ø degree position, coincident with a constraint. In this case the attenuation in the signal direction is small compared to that of Figure 1.6. But the attenuation is greater than that attained in Figure 1.5 when the signal was arriving from the Ø degree direction and close to three constraints, instead of just a single constraint.

Figure 1.8 is a plot of the converged array gain in the signal direction, for all possible signal arrival directions. This plot is obtained by placing the signal at a specified direction, calculating the optimum weight vector for this signal configuration, using this optimum

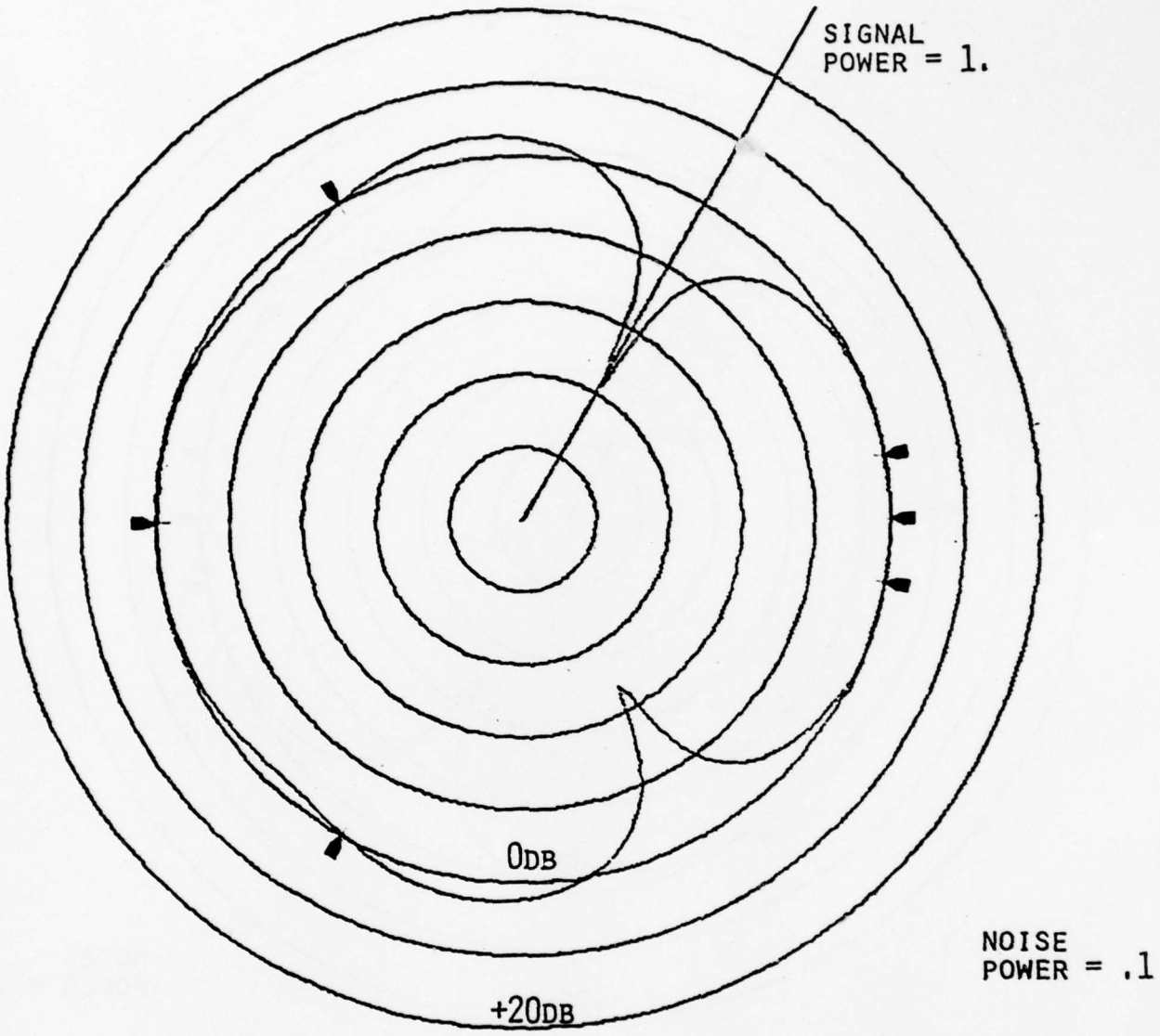


FIG. 1.6 ANTENNA ARRAY DIRECTIVITY PATTERN WITH SOFT CONSTRAINTS AND ONE SIGNAL

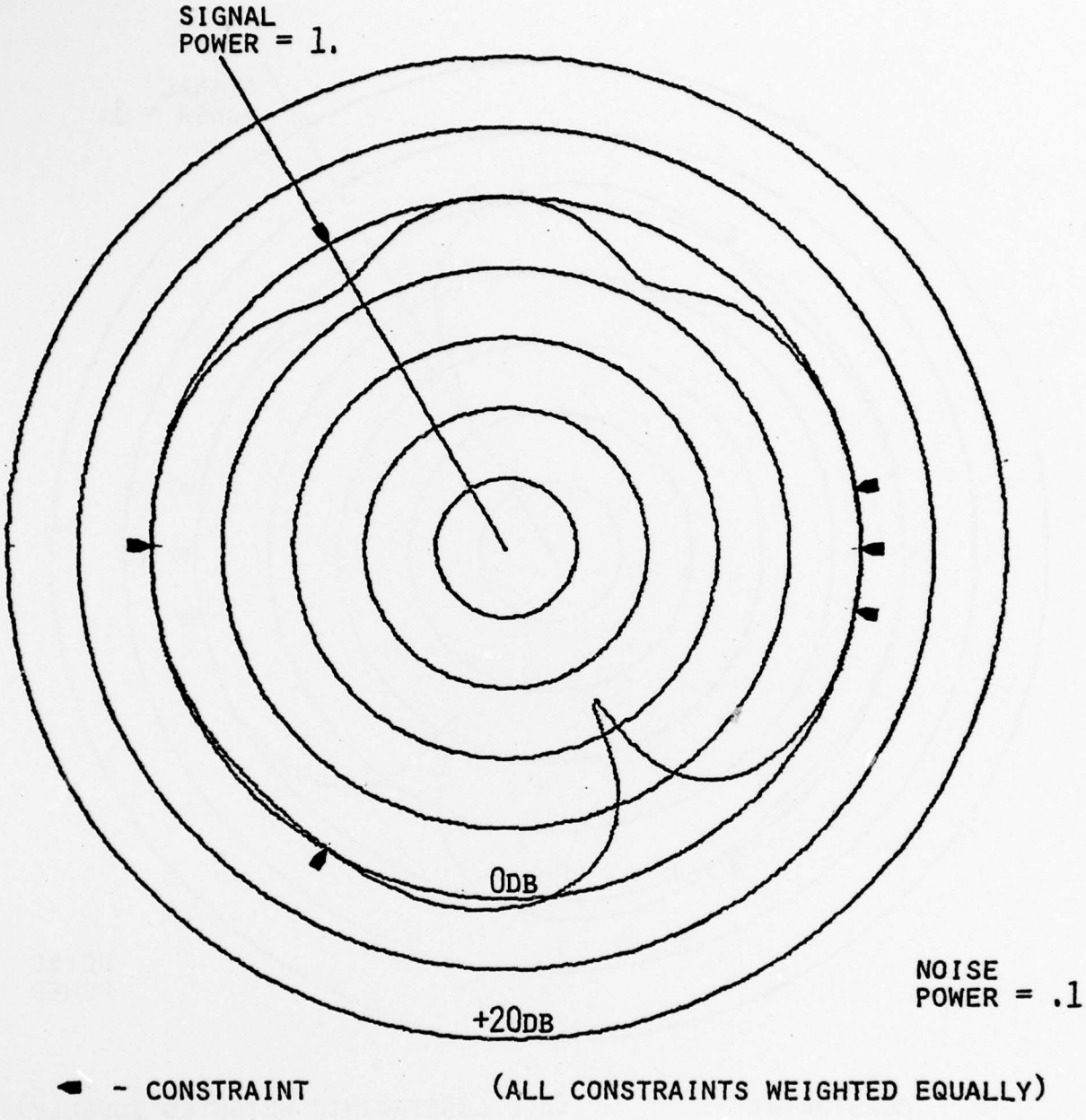
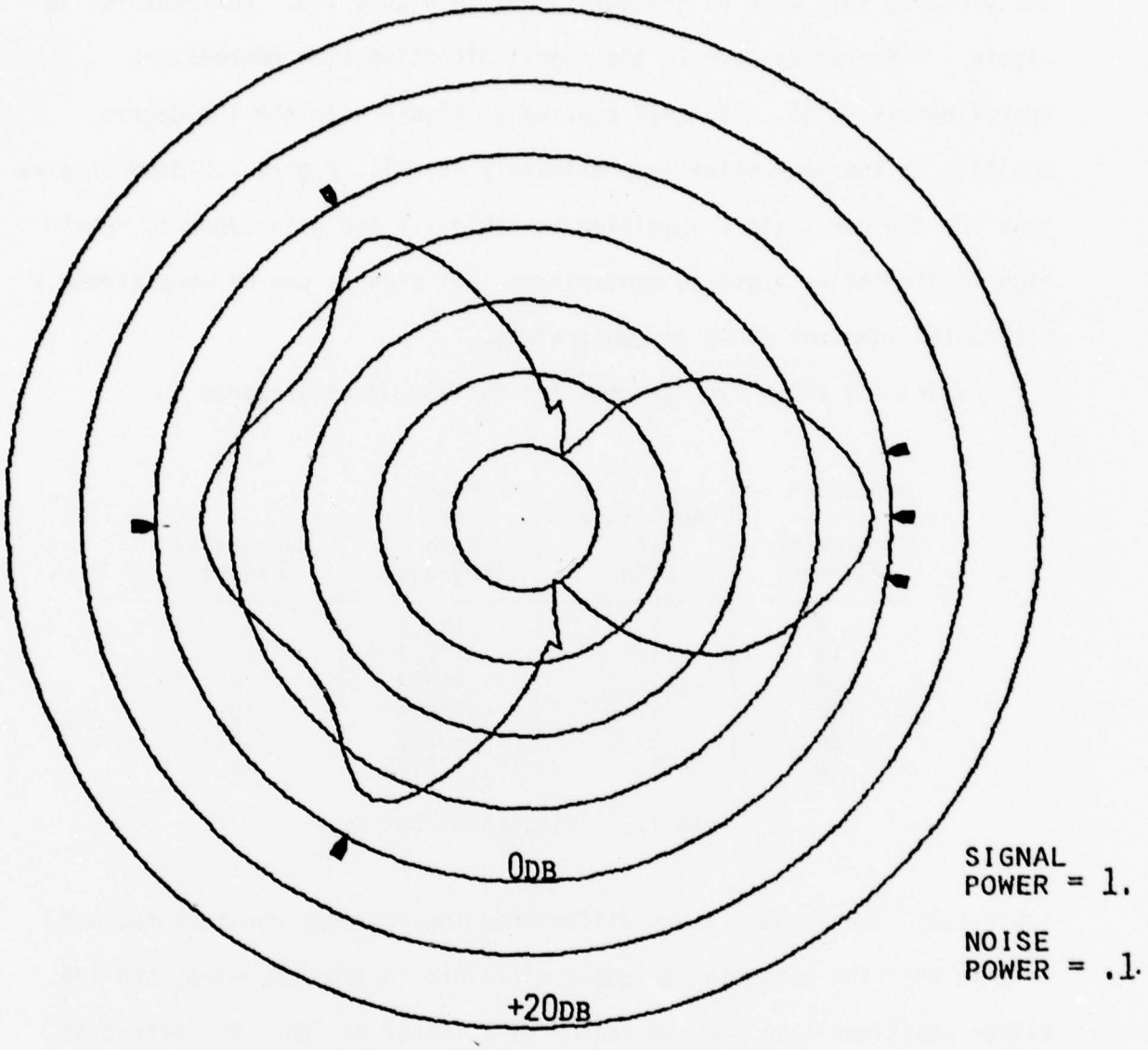


FIG. 1.7 ANTENNA ARRAY DIRECTIVITY PATTERN WITH SOFT CONSTRAINTS AND ONE SIGNAL



◄ - CONSTRAINT (ALL CONSTRAINTS WEIGHTED EQUALLY)

FIG. 1.8 ANTENNA ARRAY GAIN IN DIRECTION OF RECEIVED SIGNAL WITH SOFT CONSTRAINTS

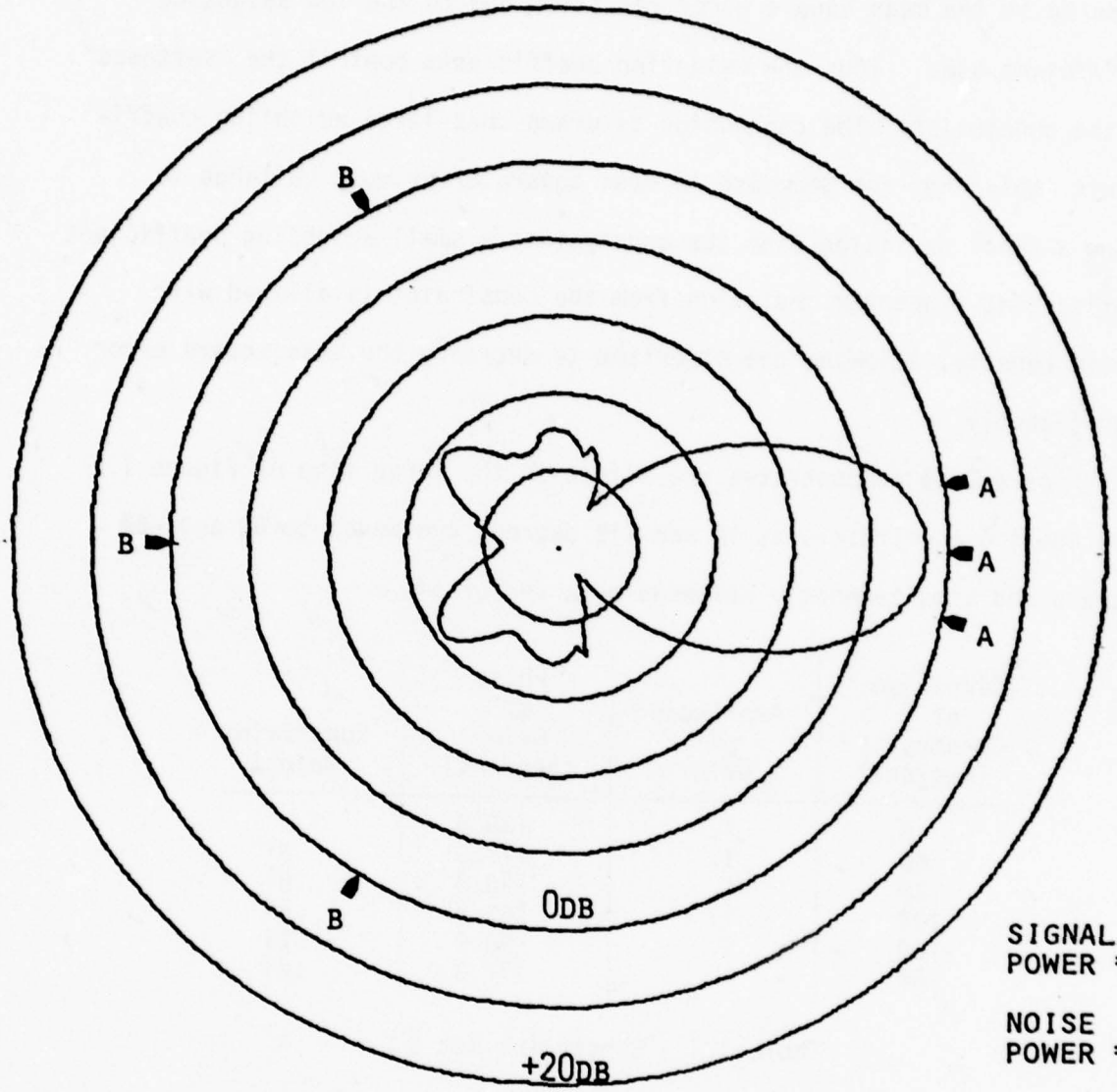
weight vector to calculate the antenna array gain in the signal direction, and plotting this gain as a single point in Figure 1.8. For example, in Figure 1.7 the array gain in the signal direction (120 degrees) is approximately -6 dB. The gain plotted on Figure 8 in the 120 degree position is the same value (approximately -6 dB). Figure 1.8 demonstrates that for the constraints specified in Table 1.1 the gain tends to remain high in directions close to constraints, but signals can be more strongly attenuated when not close to constraints.

Figure 1.9 shows similar data for the set of constraints in

Direction of Constraint (degrees)	Amplitude of Gain	Phase of Gain (degrees)	Constraint Weight
0.	1.	180.0	1.
10	1.	177.3	1.
120	1.	-90.0	.01
180	1.	-180.0	.01
-120	1.	-90.0	.01
-10	1.	177.3	1.

Table 1.2 - Constraint Set 2

Table 1.2. These constraints differ from the previous constraints, but only in that the weighting on the constraints in the 120, -120, and 180 degree positions have been decreased by a factor of 100. The effect of this decrease on the weighting of the soft constraints can be seen in Figure 1.9 as a weakening of the constraints in these directions; the array gain in the signal direction when the signal is arriving from directions close to these constraints is greatly reduced from the previous case (Figure 1.8). This arises because the mean square error can now be reduced significantly by decreasing the array gain in the



◄ - CONSTRAINT (A CONSTRAINTS WEIGHTED AT 1.,
B CONSTRAINTS WEIGHTED AT .01)

FIG. 1.9 ANTENNA ARRAY GAIN IN DIRECTION OF RECEIVED SIGNAL WITH SOFT CONSTRAINTS

signal direction. The constraint error incurred by doing this is small compared to the mean square error reduction due to the low weighting coefficient used. Thus the weighting coefficients control the "softness" of the constraint. The conclusion is drawn that large weighting coefficients imply that the decrease in mean square error must be large to allow a small deviation from the constraint; a small weighting coefficient implies that a greater deviation from the constraint is allowed with little penalty, allowing the algorithm to decrease the mean square error significantly.

Figure 1.1Ø demonstrates the effect on the large lobe of Figure 1.9 when the two constraints at 1Ø and -1Ø degrees are moved to 6Ø and -6Ø degrees and simultaneously weakened by a factor of

Direction of Constraint (degrees)	Amplitude of Gain	Phase of Gain (degrees)	Constraint Weight
Ø.	1.	18Ø.Ø	1.
6Ø	1.	177.3	.Ø1
12Ø	1.	-9Ø.Ø	.Ø1
18Ø	1.	-18Ø.Ø	.Ø1
-12Ø	1.	-9Ø.Ø	.Ø1
-6Ø	1.	177.3	.Ø1

Table 1.3 - Constraint Set 3

1ØØ. This set of constraints are presented in Table 1.3. Figures 1.9 and 1.1Ø shows that when the two strong constraints are at the 1Ø and -1Ø degree positions the angular sector over which a signal can be received without significant attenuation is much broader than when a single strong constraint is present.

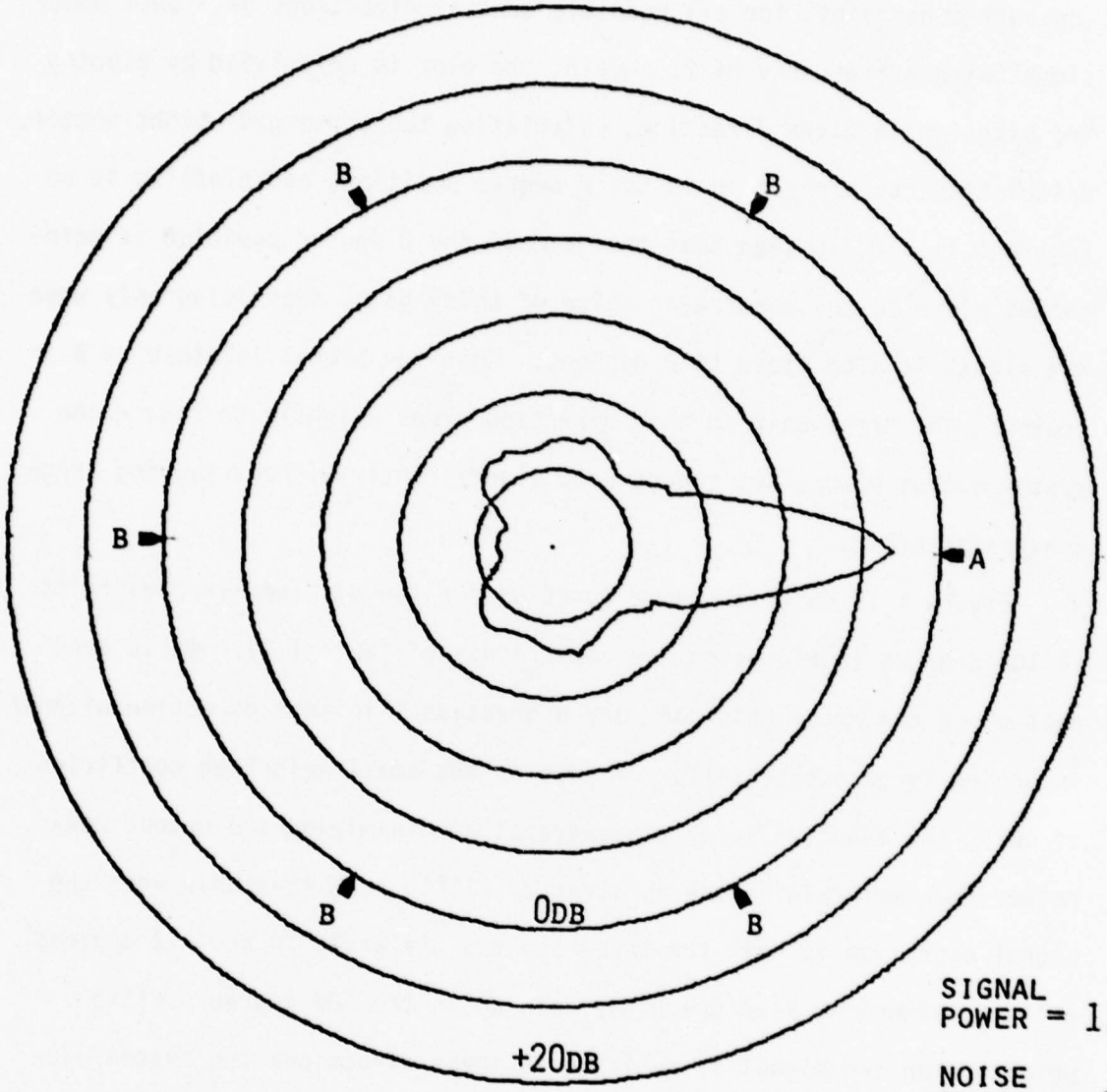


FIG. 1.10 ANTENNA ARRAY GAIN IN DIRECTION OF RECEIVED SIGNAL WITH SOFT CONSTRAINTS

Figure 1.11 shows the gain in the Ø degree direction maintained by the soft constraint, for all possible arrival directions of a unit power signal with a frequency of 2. Again, the plot is calculated by placing the signal at a given direction, calculating the converged weight vector, calculating the array gain in the Ø degree position, and plotting it on Figure 1.11. It is seen that the gain in the Ø degree position is maintained close to the constraint value of unity gain, decreasing only when the signal is also close to Ø degrees. When the signal is close to Ø degrees, the array gain in this direction drops slightly to reduce the system output power, but cannot drop significantly without causing large constraint errors.

Figure 1.12 shows the same function for the much weaker constraint at 18Ø degrees (again using the constraints of Table 1.2). It is seen that since the array gain can vary a great deal in this direction without incurring large constraint error (due to the small weighting coefficient of .Ø1), the adaptive array concentrates on minimizing the output power rather than maintaining the constraint. It is seen that only when the signal direction is near the three constraints at Ø, 1Ø and -1Ø degrees will the algorithm keep the array gain up in the 18Ø degree position, because when the signal is arriving in these directions the system output power cannot be significantly reduced by violating the constraint at 18Ø degrees. In this case the algorithm will keep the array gain close to the constraint value.

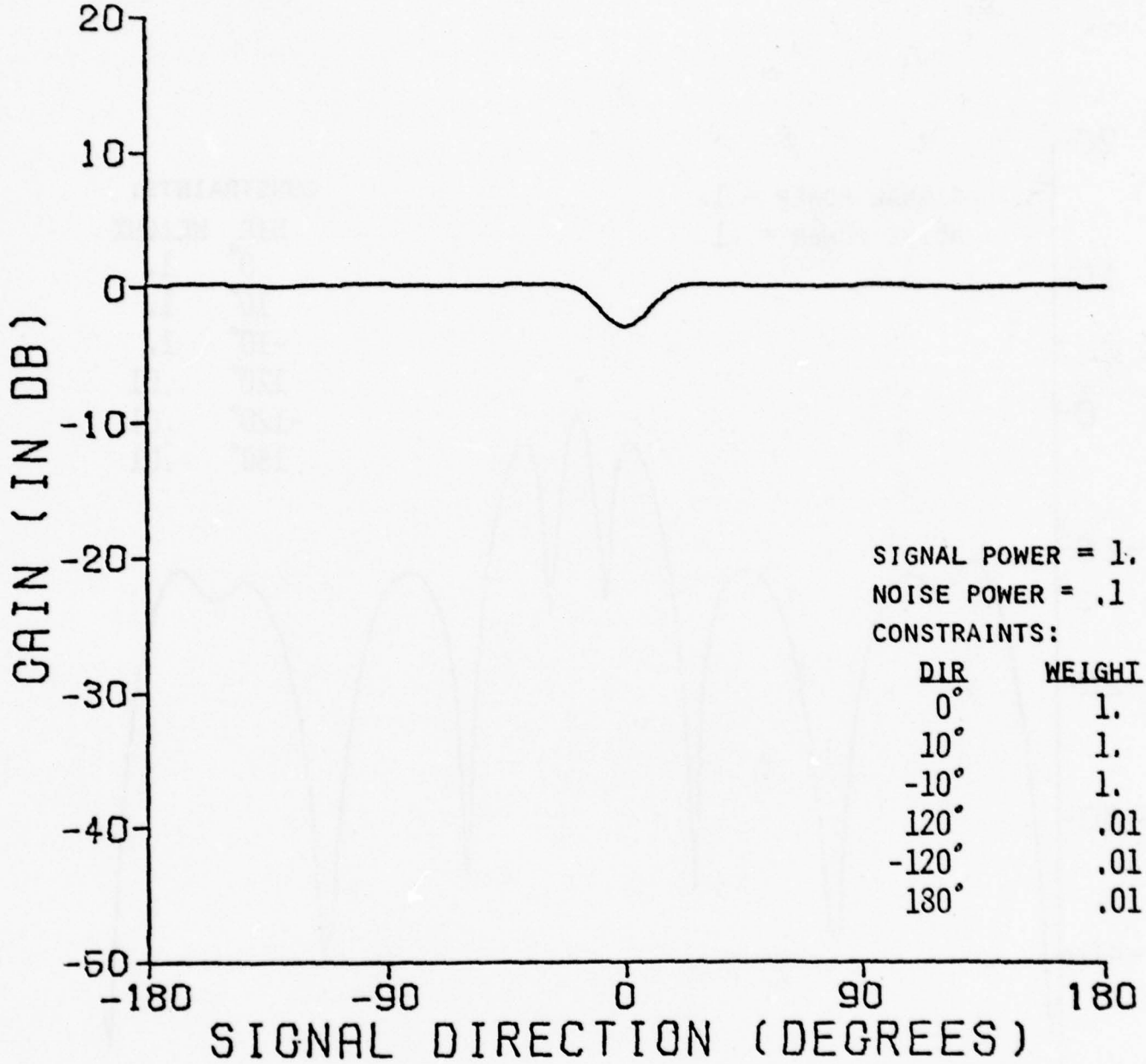


FIG. 1.11 ANTENNA ARRAY GAIN IN DIRECTION 0° WITH SOFT CONSTRAINTS AS A FUNCTION OF ARRIVAL ANGLE OF A SINGLE SIGNAL

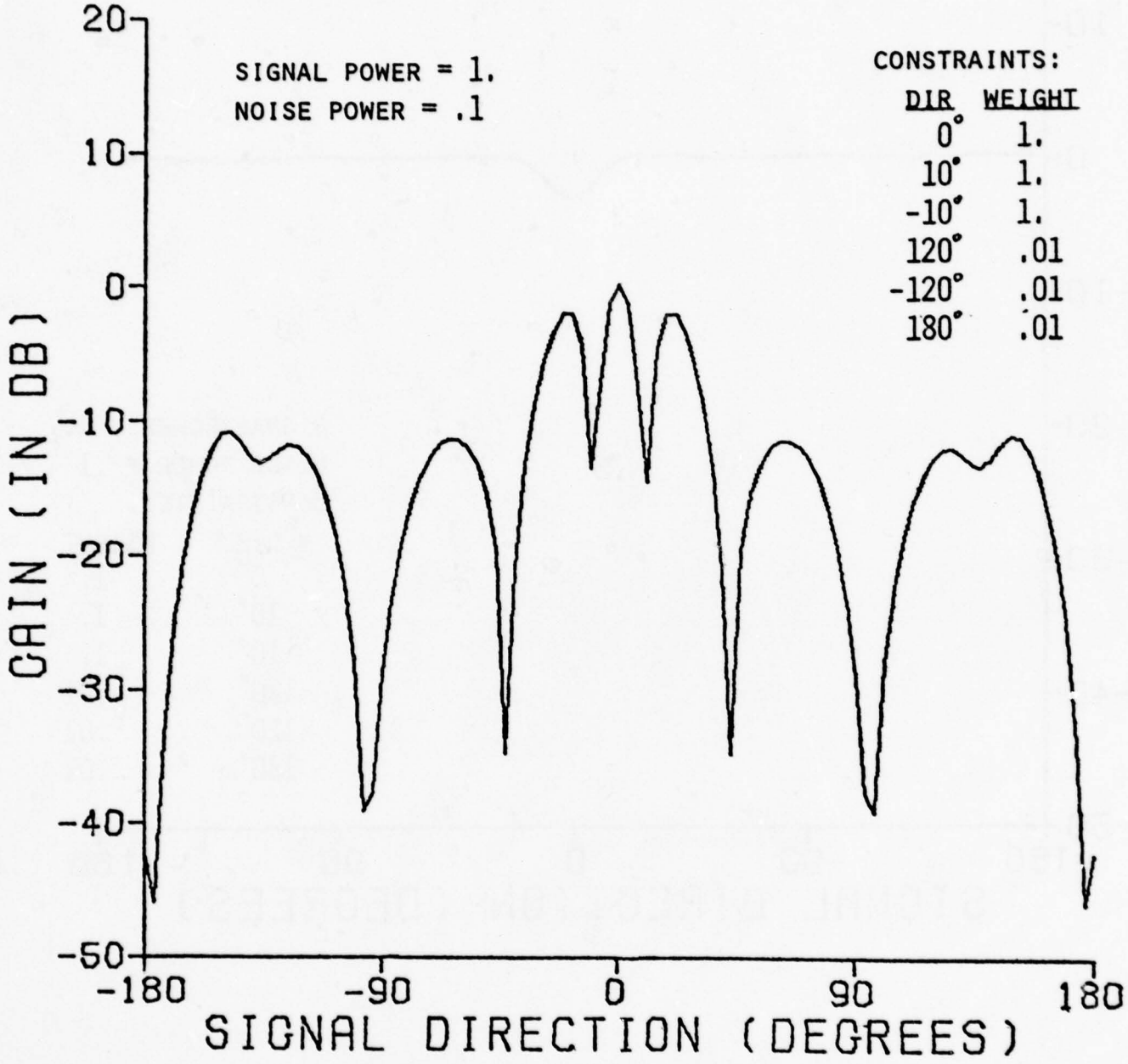


FIG. 1.12 ANTENNA ARRAY GAIN IN DIRECTION 180° WITH SOFT CONSTRAINTS AS A FUNCTION OF ARRIVAL ANGLE OF A SINGLE SIGNAL

PART 2. AN IMPROVED ADAPTIVE POWER SEPARATOR

2.1 General Utility of a Power Separator

All radio communication systems require a filter to separate the desired information signal from background noise. Since the background noise also contains undesired information signals as well as broadband random signals, the receiver must have some information about the characteristics of the desired signal. In simple radio receivers the frequency of the signal, as well as the modulation technique, is all that is specified. However, in situations where there are many signals in a small frequency band it may be difficult, or impossible, to discriminate by frequency alone. In some cases the task of separating the desired signal can be simplified by using a spatial filter. The spatial filter may be a high gain antenna or a high gain antenna array. Most of the work on adaptive antenna systems has concentrated on array systems [8,9,19]. (This is because it is easier to delay a signal than it is to move the hardware. Also, multiple electrical circuits can be used to create independent antenna patterns from a single physical array.)

Another technique for identifying the desired signal requires knowledge of its time characteristics. The receiver correlates the input with a known reference sequence [20]. This effectively eliminates all signals which are not correlated with the desired signal. Thus with the proper choice of a reference signal, only the desired signal will correlate with the reference and hence only the desired signal will appear at the receiver output.

All the techniques for identifying the desired signal use some knowledge -- frequency, direction, or correlation -- of the desired signal to separate the desired signal from the background noise. While the use of adaptive systems does not fundamentally change communication system design, adaptive systems do have two significant advantages: 1) nearly optimal processing of signals, and 2) the ability to accommodate deviations from the expected signal characteristics. Thus, if used properly, adaptive systems reduce the amount of knowledge needed to separate a desired signal from other signals and noise.

John Treichler showed [11,21] that an adaptive system known as the leaky adaptive line enhancer (LALE) could be used to discriminate between high and low power signals. A natural application is anti-jam protection against high power jamming signals. However, it is not reasonable to view the application of power discrimination as limited to protection from jamming signals. Power discrimination should be considered as a general technique which can be used by systems to separate a desired signal from other signals; and, as such, power separation systems take a place alongside frequency sensitive systems, direction of arrival sensitive systems and correlation sensitive systems.

An example of an application for a power separator is provided by the work of Chestek [22] (also see the previous section of this report). Chestek has developed an adaptive antenna array which normally has an omnidirectional pattern. When a signal impinges on the array, an adaptive signal processor tries to cancel the signal by forming a null in the direction of the signal. However, power discrimination is built into the algorithm so that weak signals are only slightly attenuated but strong

signals are greatly attenuated. This signal power discrimination (see also Zahm [14]) results in an inversion of signal powers. That is, the strongest input signals to the array processor become the weakest output signals. The advantage of this system is that it passes the desired signal even when the frequency and direction of the desired signal are unknown. An array constrained to look in a given direction [2] may not perform well if the direction of the signal is imperfectly known. Similarly, if the frequency of the desired signal is imperfectly known, because of doppler shifts for example, a fixed frequency system may work poorly.

2.2 Two Special Applications for a Power Separator

A specific application of power separation is the alleviation of dynamic range (ratio of largest signal to smallest signal) problems. The first dynamic range problem is associated with the use of the Widrow-Hoff least mean square (LMS) algorithm [23,24] to adapt filters when the input consists of signals with greatly differing powers. The second dynamic range problem concerns the cost of building hardware with a large dynamic range.

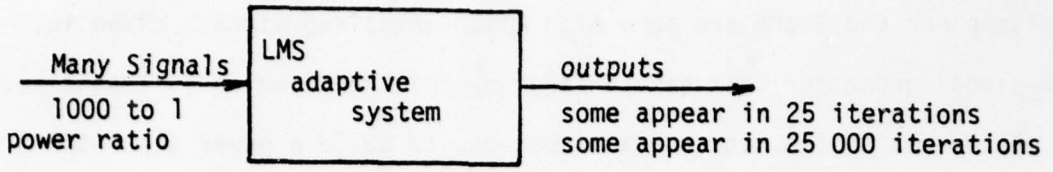
Signals with widely differing power levels can cause convergence problems for filters which use the LMS algorithm for adaption. Work by Widrow [10,11] has shown that the convergence of an LMS adaptive filter consists of many exponential modes. Each mode is associated with a particular input signal, and the time constant of each mode is inversely proportional to the power of the mode. (Since the eigenvalues of the autocorrelation matrix reflect the powers of the modes, this convergence problem is sometimes called the eigenvalue spread problem.) A problem

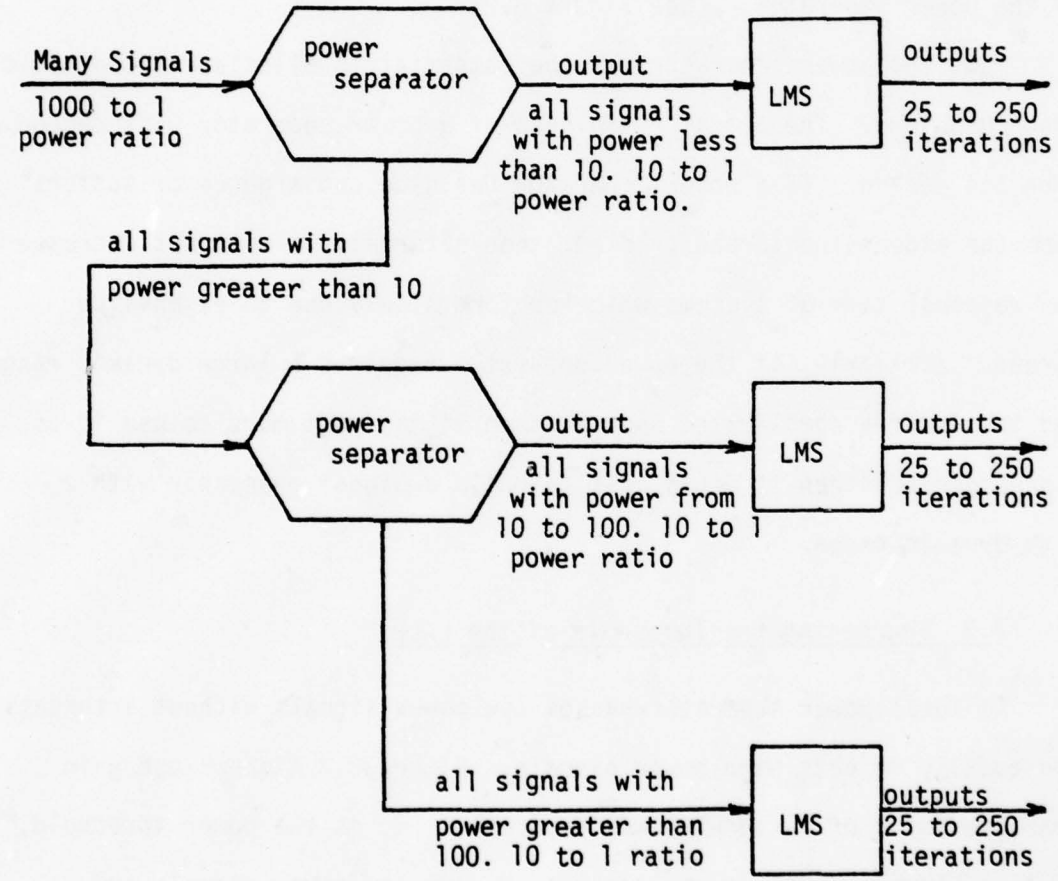
occurs when there are several signals with differing power levels. Because of the convergence properties of an LMS filter, the high power modes will converge faster than the low power modes. By varying an adaptive parameter, μ, the convergence rates can be altered, but there is a limit to the rate of convergence. If $\mu > 1/\lambda_{max}$, where λ_{max} is the power of the most powerful input, then the filter will be unstable. Thus $\mu < 1/\lambda_{max}$ for stability. The time constant for the fastest mode is:

$$\tau_{max} > \frac{1}{4\mu\lambda_{max}} .$$

But if the system is stable this must be greater than 1/4. Usually for good filter performance $\mu < 1/100\lambda_{max}$ which implies that $\tau > 25$ for the high power mode. Any low power mode will converge more slowly. In fact, $\tau_{low} = \frac{\lambda_{max}}{\lambda_{low}} \tau_{max}$, so, if there is a 1000 to 1 power ratio, the low power mode will take 1000 times longer to converge than the high power mode. This may be an unacceptably long time. By using power separators to decompose the input signal into groups of component signals with only a small power spread in each group, a separate adaptive filter can be used for each group and optimized for the power levels in that group. Since each group of component signals has a small power spread, the time constants will have only a small spread. (See figure 2.1)

The second dynamic range problem is associated with the cost of hardware. All hardware has a limited dynamic range. If too large a signal is applied to the system, the system will saturate with many undesirable side effects. If too small a signal is applied to a system, there will be no output apart from noise. The dynamic range of a system must accommodate all signals of interest plus any higher power signals





Power Separators used to alleviate the eigenvalue problem

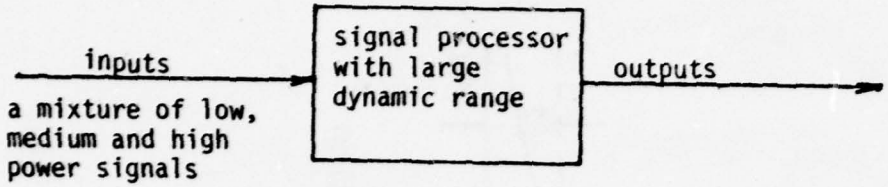
Figure 2.1

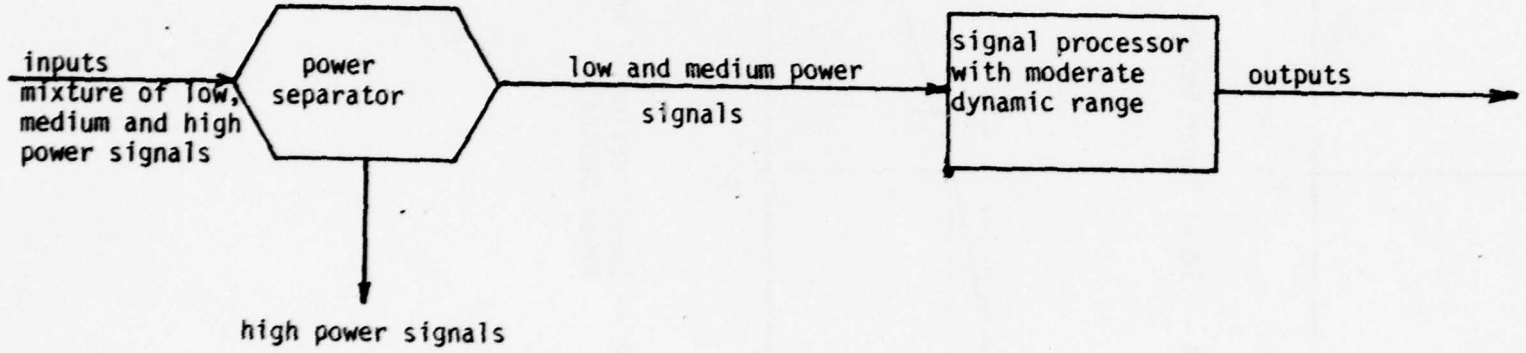
that are in the input. Thus if the signals of interest are low to medium power and there are some high power undesired signals mixed in, the signal processor must have a large dynamic range which is expensive. It may be reasonable, in such a situation, to build a power separator with high dynamic range. The output of the power separator would be the low and medium power signals, which could then be applied to a signal processor that has only a moderate dynamic range. The idea is economical if the reduced cost of the signal processor offsets the additional cost of the power separator. (See Figure 2.2)

Thus the power separator has the potential to alleviate two dynamic range problems. The actual usefulness of a power separator will depend upon its design. If a power separator has slow convergence or suffers from the eigenvalue spread problem, then naturally it will not decrease the response time of systems which perform slowly due to eigenvalue spread. Similarly, if the power separator requires a large dynamic range and is itself a complicated system, then it may cost more to use it as a preprocessor than it would cost to build a signal processor with a high dynamic range.

2.3 Sharpening the Threshold of the LALE

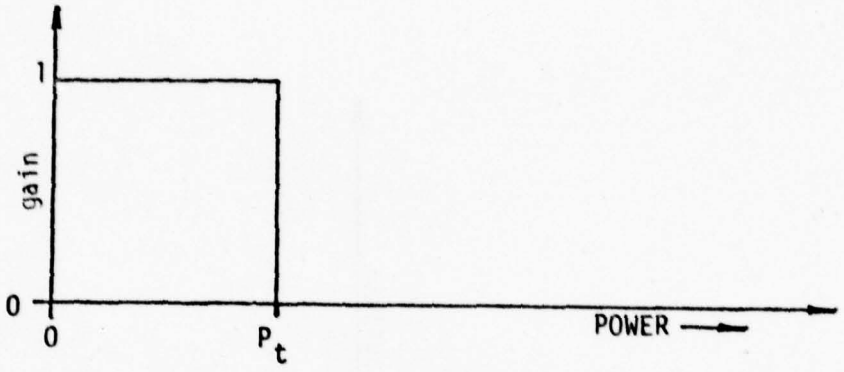
An ideal power separator passes low power signals without attenuation and totally rejects high power signals. Figure 2.3.A shows the gain characteristic of an ideal power separator. P_t is the power threshold, so those signals with power less than P_t are low power signals and therefore passed with a gain of 1, while those signals with a power greater than P_t are high power signals and therefore eliminated.



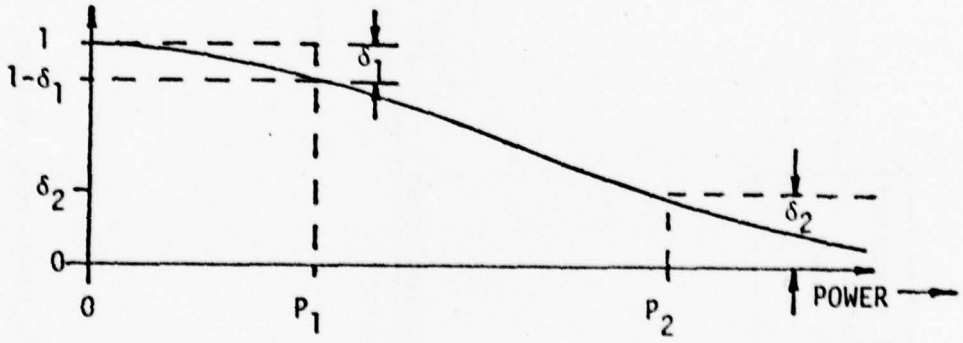


Use of a Power Separator to reduce the dynamic range requirements of a signal processor

Figure 2.2



Part A -- Ideal Power Separator



Part B -- Specification of bounds for a practical Power Separator

Figure 2.3

The leaky adaptive line enhancer (LALE) has a gain versus input power curve which can be used for power separation. Figure 2.4 shows a block diagram of a LALE, while figure 2.5 shows the gain curves for the LALE. The curve labeled Y out in Figure 2.5 corresponds to the gain from the input to the Y output. This output can be used to pass high power signals and to reject, or at least greatly attenuate, low power signals. The curve labeled E_{out} in Figure 2.5 corresponds to the gain from the input to the E output of the LALE. The E output can be used to pass low power signals and to reject high power signals.

The transitions from high gain to low gain are not very sharp. The ideal gain characteristic for a power separator would have a gain of one when the input power was less than a predetermined threshold, and zero if the input power was greater than the threshold, as shown in Figure 2.3.A. Realistically we cannot expect to achieve a discontinuity in the gain characteristic of a power separator. However, it is not unreasonable to require that:

a) $1-\delta_1 < g < 1$, $0 < P_{in} < P_1$

b) $g < 1$, $P_1 < P_{in} < P_2$

c) $g < \delta_2$, $P_2 < P_{in}$

where g = filter gain at signal frequency

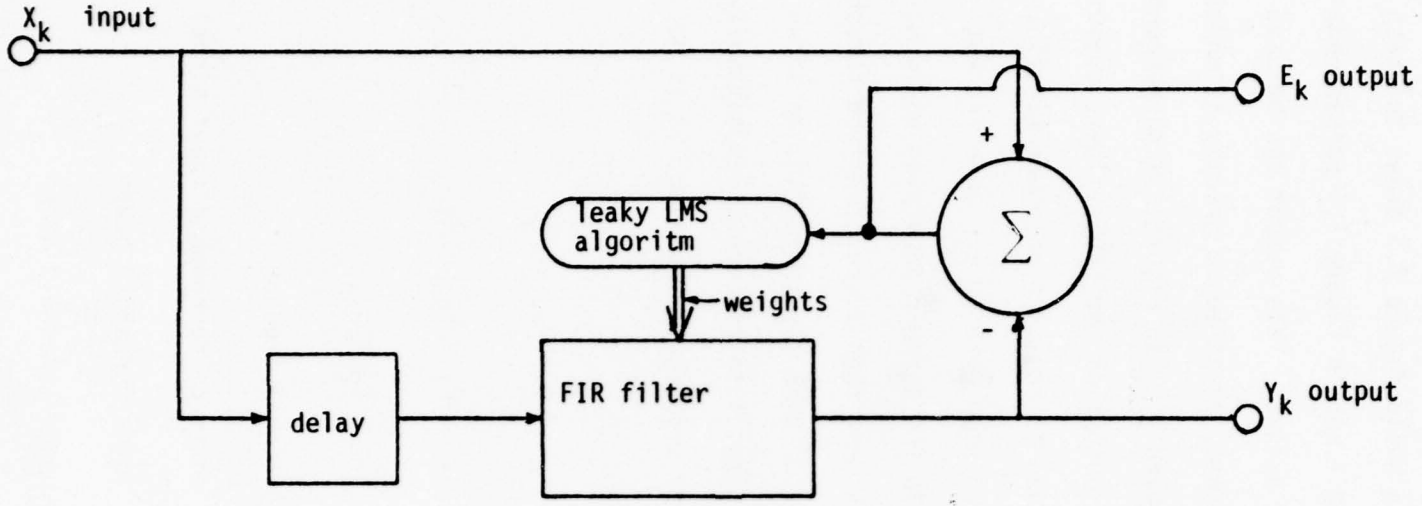
P_{in} = power of narrowband input signal

P_1 = lower threshold power

P_2 = upper threshold power

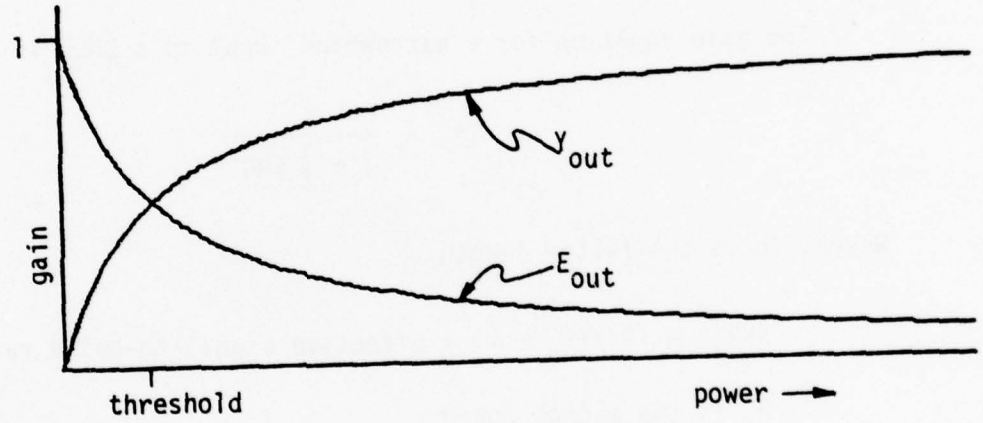
δ_1 = maximum allowable gain loss in pass region

δ_2 = maximum allowable gain in rejection region



Leaky adaptive (spectral) line enhancer (LALE)

Figure 2.4



Gain curves for LALE

Figure 2.5

(See Figure 2.3.B for example.)

By reducing δ_1, δ_2 and P_2-P_1, we can make the filter approach the ideal characteristic, and in fact, if δ_1=0, δ_2=0 and P_2=P_1 then the characteristic is ideal.

Obviously we could define a similar characteristic for a filter that rejected low power signals and passed high power signals (in fact g' = 1-g is such a characteristic), but we will not treat the case of the low power rejection filter because it is so easily derived from the high power rejection filter. Therefore, whenever a power separator is mentioned it will mean a filter which passes low power signals and rejects high power signals.

The gain equation for a narrowband input to a LALE is [11,21]:

$$g = \frac{1}{1 + \frac{n}{2} SNR'} \quad . \qquad 2.1$$

Where: n is the filter length

$SNR' = P_s/(P_n + \gamma)$ = effective signal-to-noise ratio

P_s is the signal power

P_n is the noise power

γ is a parameter of the LALE .

This curve is monotone decreasing since:

$$\frac{d\,g}{d\,P_s} = \frac{-\frac{n}{2(P_n + \gamma)}}{\left(1 + \frac{n}{2} SNR'\right)^2} < 0 \quad . \qquad 2.2$$

Where n and $\gamma > 0$. The selection of optimal n and γ, both filter design parameters, will cause the gain at each threshold to just meet the gain specification at the threshold, i.e. a gain of $1-\delta_1$ at P_1 and δ_2 at P_2. However, only one of the specifications (δ_1,P_1) or (δ_2,P_2) is needed to select the ratio $n/(P_n + \gamma)$ which is given by:

$$\frac{n}{P_n + \gamma} = \frac{2}{P_1} \frac{\delta_1}{1-\delta_1} = \frac{2}{P_2} \frac{1-\delta_2}{\delta_2} \quad . \qquad 2.3$$

As a result the four quantities, δ_1, P_1, δ_2, P_2 cannot be arbitrarily selected. Any three specify the fourth.

Equation 2.3 can be rewritten as:

$$W_t = \frac{P_2}{P_1} = \frac{1-\delta_1}{\delta_1} \frac{1-\delta_2}{\delta_2} \quad . \qquad 2.4$$

In this form, the width of the transition region, W_t, is related to the quality of the pass and reject regions where the quality of the pass region is better for smaller values of δ_1 and the quality of the rejection region is better for smaller values of δ_2. As the quality of the pass and rejection regions improves, that is as δ_1 and δ_2 tend to zero, the transition width increases.

Ideally, the designer should be able to control all four parameters, δ_1, P_1, δ_2, P_2 by adjusting design parameters. The goal of this research is to develop a power separator which gives the designer freedom to select δ_1, P_1, δ_2, and P_2. The quality of the power separator will be loosely related to the transition width. For a given δ_1 and δ_2, the better the power separator the smaller the value of W_t. However, this is

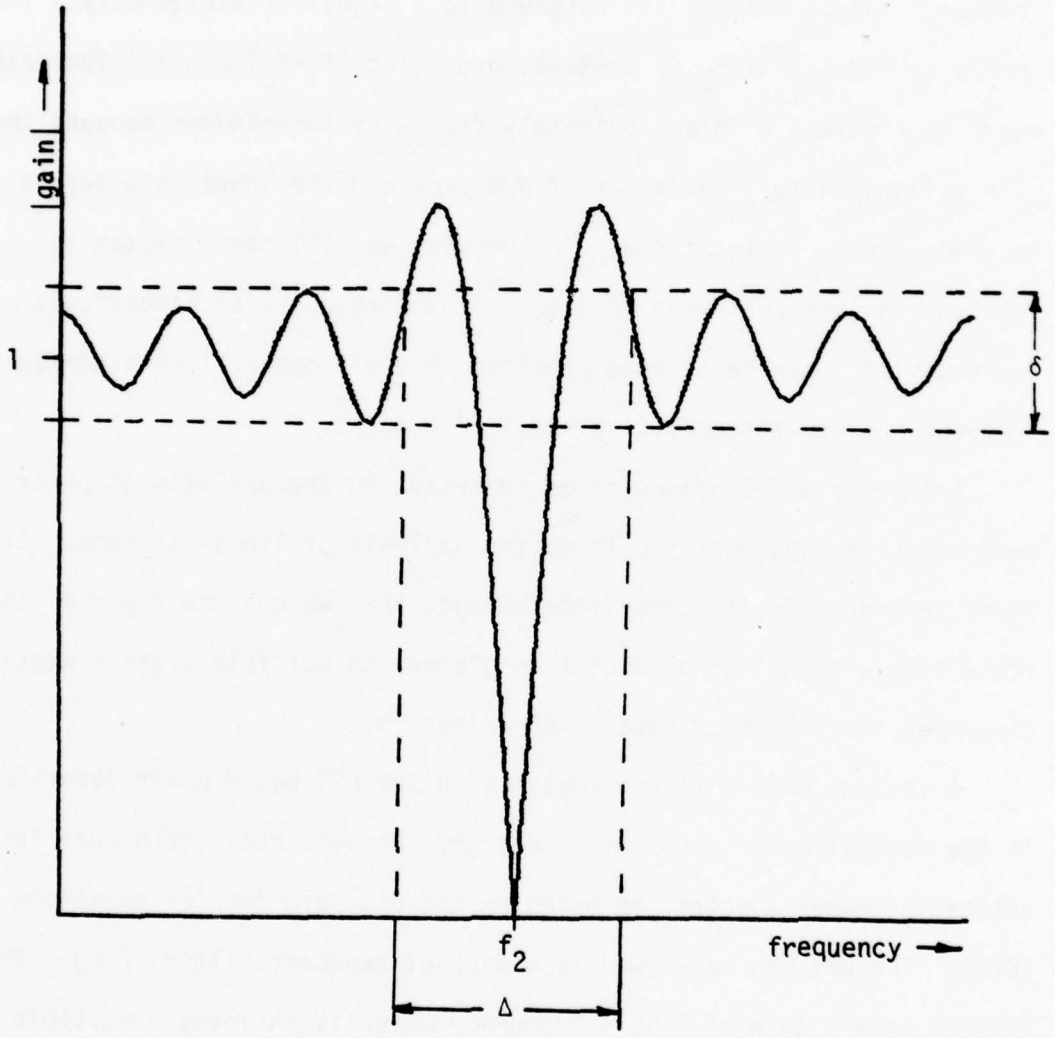
not the only measure of quality for a power separator: other factors, such as frequency independence, ease of implementation, response times and transition region characteristics must also be considered.

2.4 A DFT Based Power Separator

An obvious approach to designing a power separator is to build a system which evaluates the power spectrum of the input signal. The power spectrum is then used as an input to a system which designs a filter with zero gain at the frequencies where the power spectrum exceeds a pre-determined threshold, and a gain of one elsewhere. The input signal is passed through the filter to give an output from which the high power signals have been excluded. A power separator based on this method will have an ideal characteristic provided the power spectrum is perfectly determined and a perfect bandstop filter can be designed.

Unfortunately, perfect bandstop filters cannot be made. However, arbitrarily good filters are possible. That is no matter how small a value of ε is selected, the gain of the filter in the rejection region can always be made less than ε. (Practically, however, the accuracies of the devices used to implement the filter will limit the minimum attainable gain.) Also, the effects on neighboring frequencies can be made small. This last point is important, and it is called frequency independence.

To illustrate frequency independence, consider a power separating system with two input signals, S_1 and S_2, with powers P_1 and P_2, and frequencies f_1 and f_2 respectively. Ideally if $P_1 <$ power rejection threshold $< P_2$ then signal S_1 will be passed and signal S_2 will be rejected. However, if the bandstop filter is not ideal this may not be true. Figure 2.6 shows the transfer function of a non-ideal (but practical)



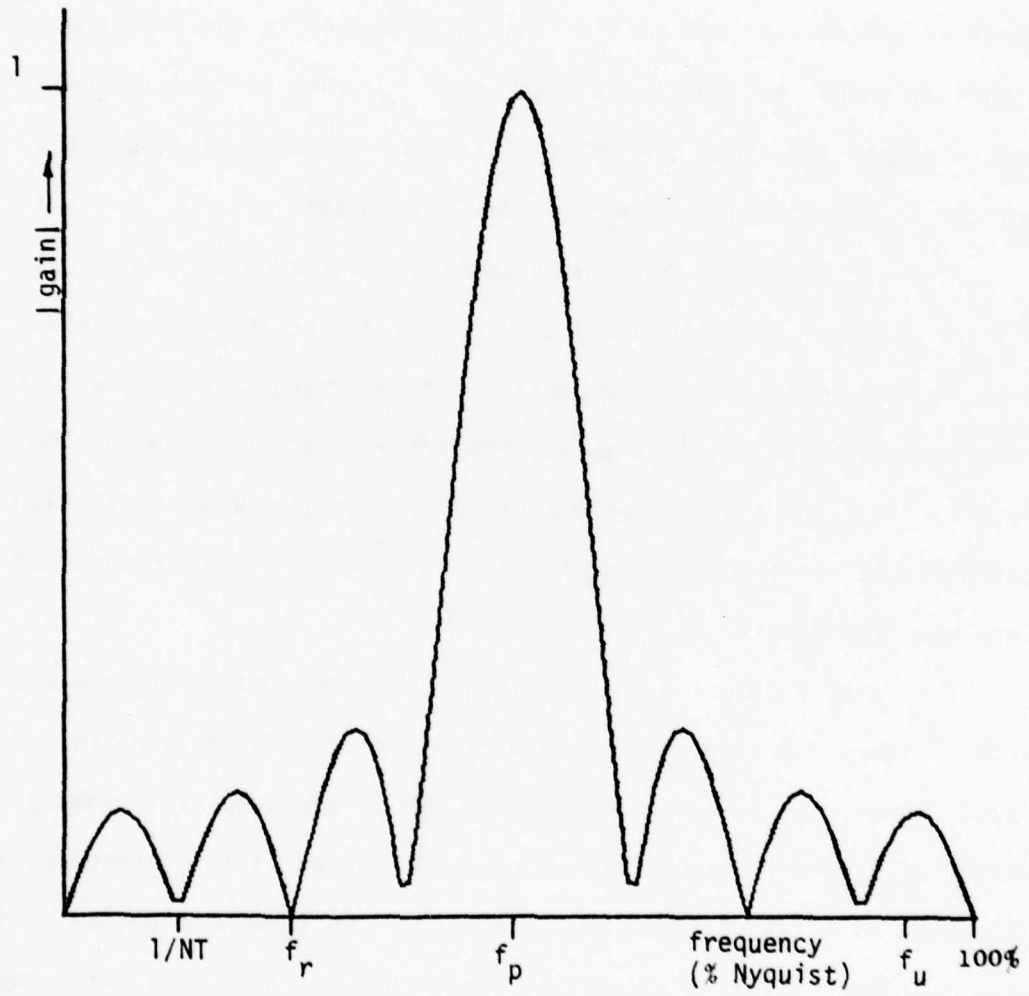
An approximately frequency independent bandstop filter

Figure 2.6

bandstop filter. It is easy to see that if f_1 is close enough to f_2, then S_1 will also be partially rejected. In this case the system is not frequency independent because the response to S_1 is dependent upon the frequency of both S_1 and S_2. Unfortunately, all of the systems we will work with are frequency dependent. A system can be called approximately frequency independent if the response to a signal S_1 varies only a little when a second signal S_2 is present, providing $|f_1-f_2| > \Delta/2$. The gain curve of Figure 2.6 is approximately frequency independent because the gain is bounded to a variation of $\delta/2$ provided the inputs are separated in frequency by at least $\Delta/2$. For brevity we will say a system is frequency independent when we mean that the response at a particular frequency, f_0, can be bounded provided that all other signals differ in frequency by at least $\Delta/2$, i.e. $|f_1 - f_0| > \Delta/2$.

Frequency independence is as important to the analysis of power separators as superposition is to the analysis of linear systems. If a power separator is frequency independent, then we can use the analysis for a single signal to predict the response to multiple signals provided they meet the frequency separation criterion.

A second, more serious problem with the DFT based power separator is the estimation of the power spectrum. Conventional techniques for estimating power spectra are based on the discrete fourier transform (DFT). The DFT can be viewed as a bank of bandpass filters [26]. These filters have a gain of 1 at a frequency which is an integer multiple of $1/(nT)$ (n is the length of the DFT, T is the sampling interval), and a gain of zero at the remaining multiples of $1/(nT)$. Figure 2.7 shows the frequency response of one such typical bandpass filter. Briefly, if a



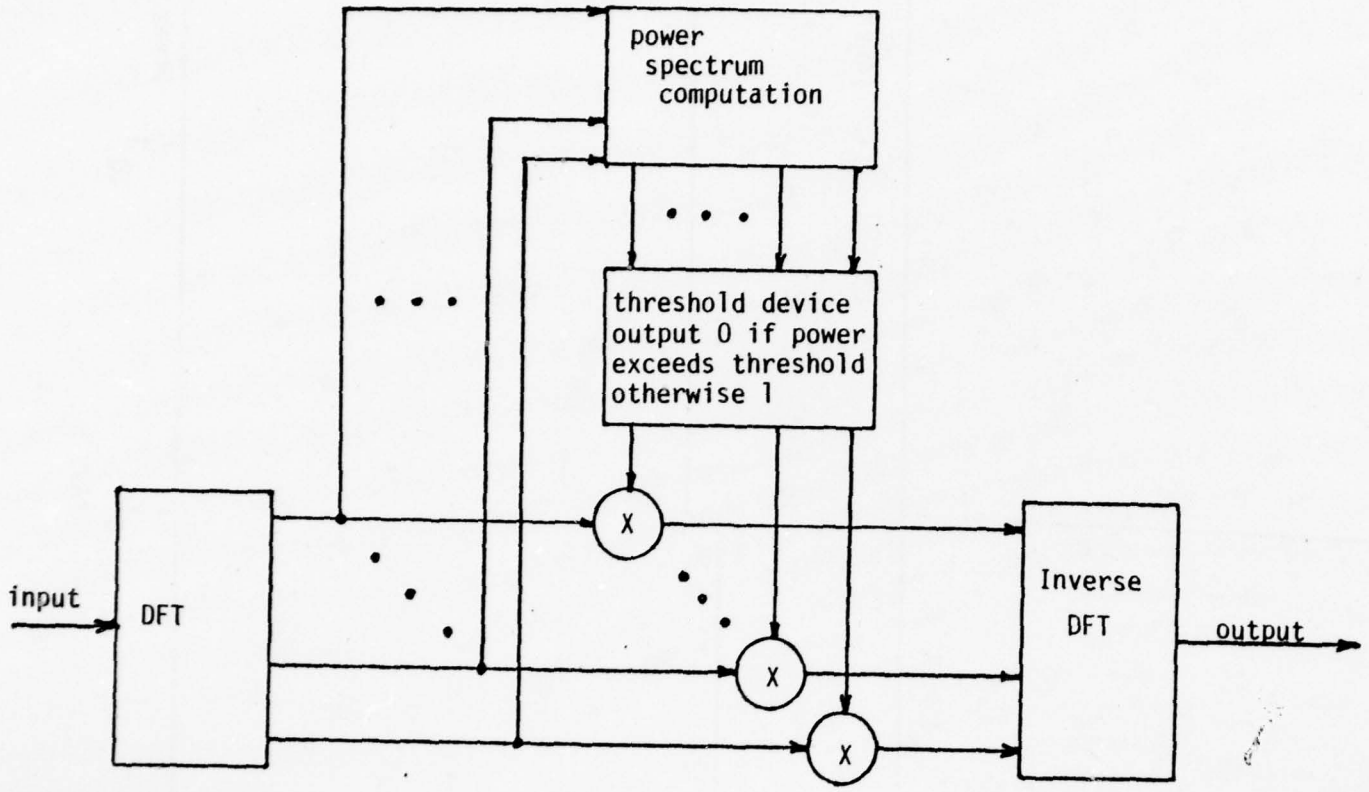
Typical bandpass filter used in the DFT

Figure 2.7

signal is at frequency f_p it will be passed with a gain of 1. Thus by examining the output power of the filter the input power can also be determined. A signal at one of the zeros, say f_r, will have no effect on the output. Hence the output of this bandpass filter reflects the amount of power in the region near f_p. However, a signal that lies in one of the sidelobes, say at f_u, will also contribute to the output power of the bandpass filter. A very high power signal at f_u can cause as much output power as a weaker signal at f_p. This will result in the erroneous assumption that there is significant power near f_p, when in fact the only input power is near f_u. Thus the DFT violates the frequency independence principle when used as a power spectral estimator.

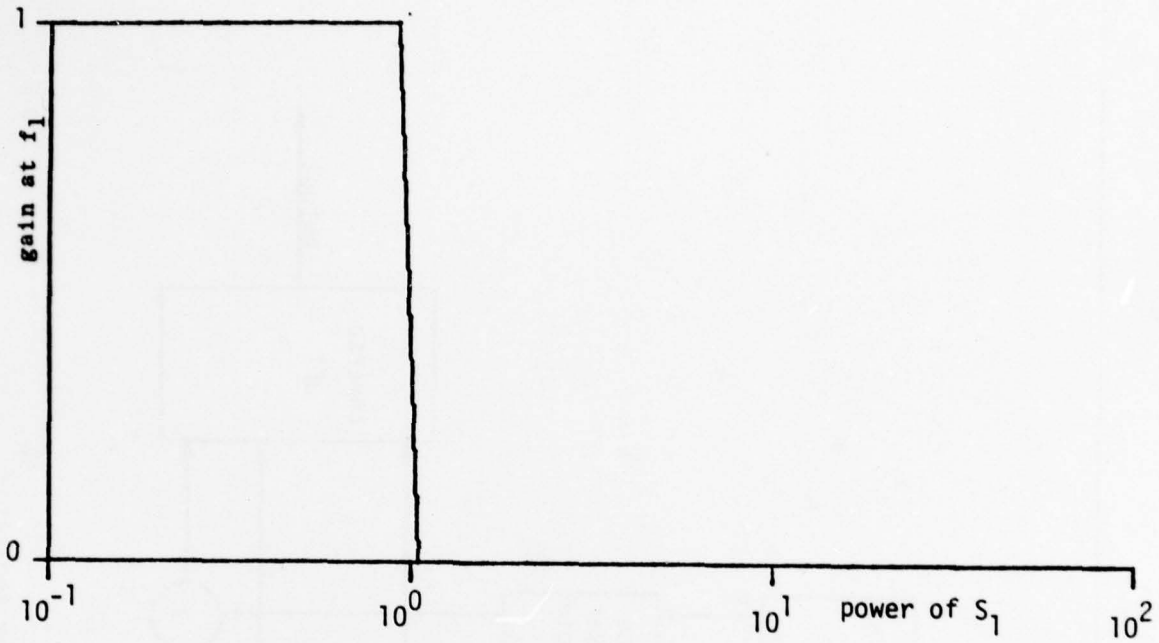
Figure 2.8 shows a block diagram of a power separator which uses a DFT for spectral estimation and a threshold device to decide whether to pass individual spectral components. The resulting spectrum, with some components eliminated, is then inverse transformed to derive a time series output. Thus the signal is transformed so that the power separation is performed in the frequency domain and the result is then transformed back into the time domain.

The results of using such a system are presented in Figures 2.9 and 2.10. Figure 2.9 displays the results of processing two signals. The first signal, S_1, is at a frequency of 25% of the Nyquist frequency. Part a of Figure 2.9 shows the gain of the system at this frequency as the power of S_1 is varied from 0.1 to 100. The threshold device activates when the power spectrum exceeds 1. Note that the gain is 1 until the power of S_1 exceeds the threshold, whereupon the gain drops to 0. Part b of Figure 2.9 shows the system response to a second signal, S_2, which has

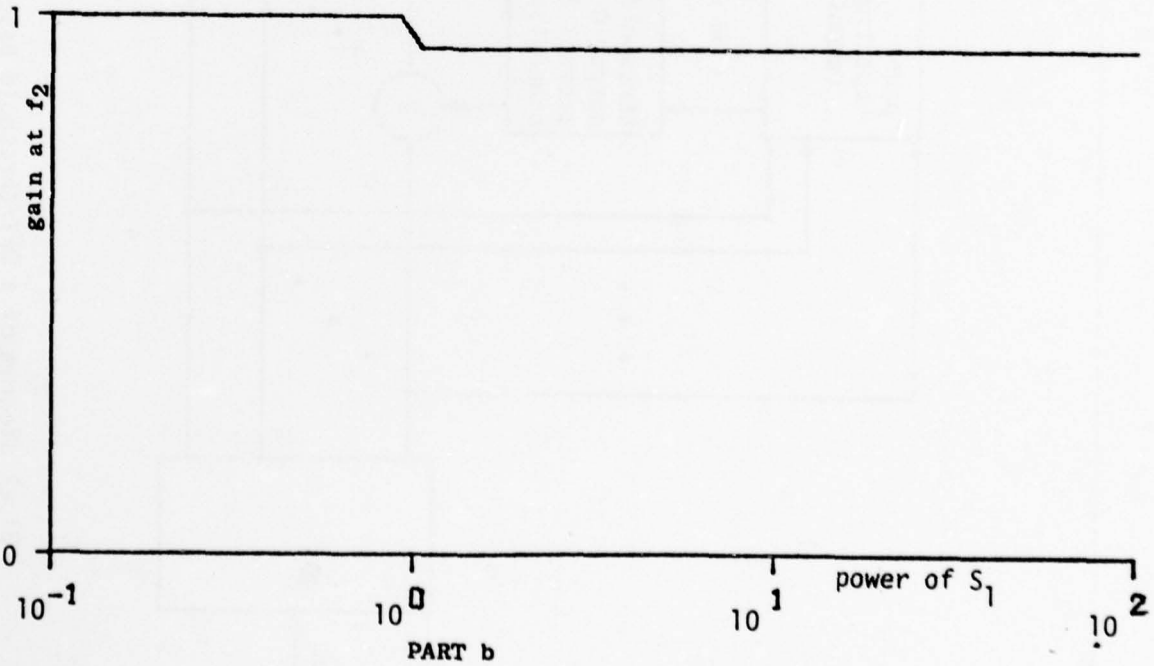


Block diagram of a DFT/threshold based power separator

Figure 2.8



PART a



PART b

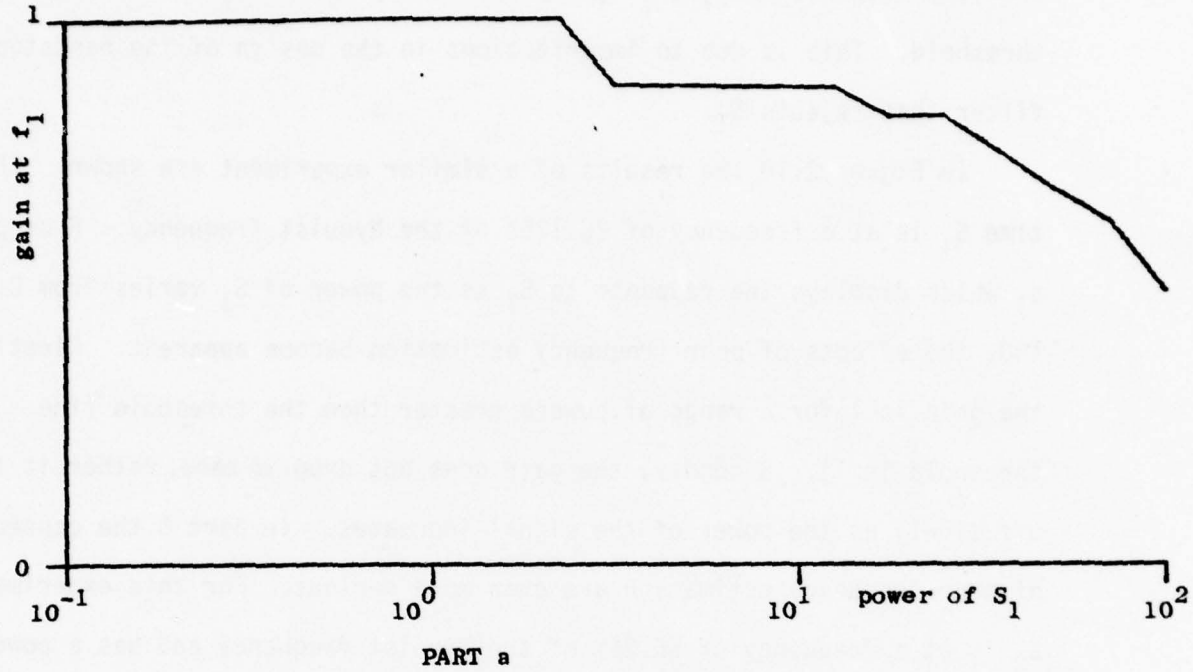
Experiment 1 using the DFT/threshold based power separator

Figure 2.9

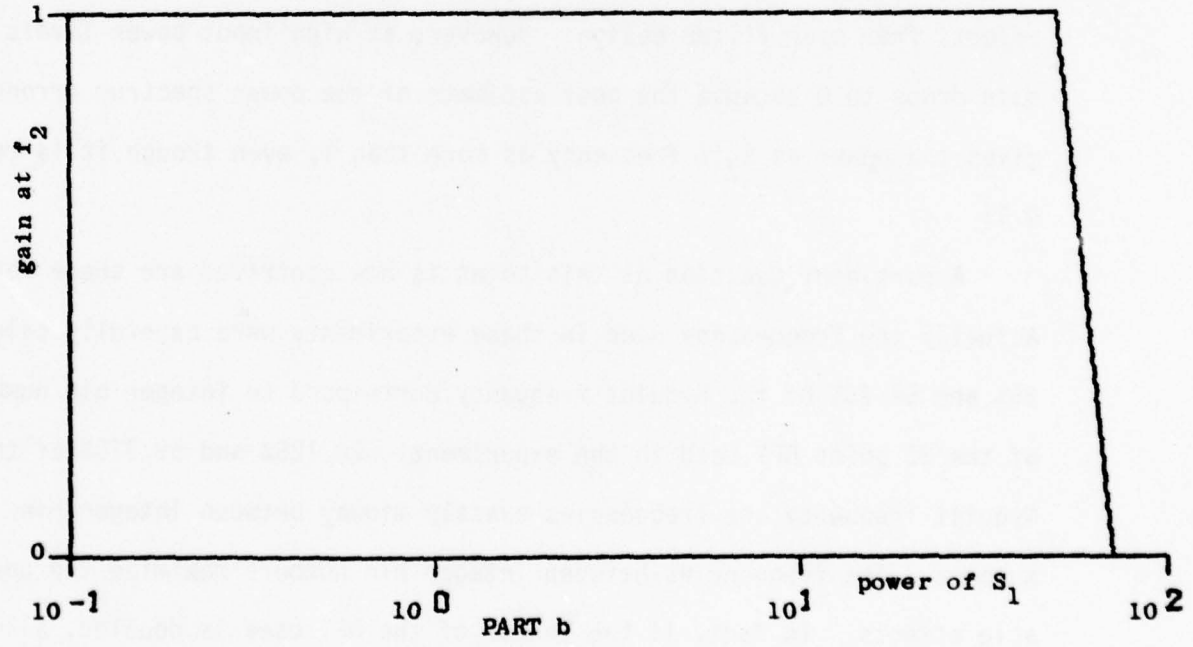
a power of 0.1 at a frequency of 59.375% of the Nyquist frequency. Notice that the response changes slightly when the power of S_1 crosses the threshold. This is due to imperfections in the design of the bandstop filter that rejects S_1.

In Figure 2.10 the results of a similar experiment are shown. This time S_1 is at a frequency of 28.125% of the Nyquist frequency. From part a, which displays the response to S_1 as the power of S_1 varies from 0.1 to 100, the effects of poor frequency estimation become apparent. Firstly, the gain is 1 for a range of powers greater than the threshold (the threshold is 1). Secondly, the gain does not drop to zero, rather it falls off slowly as the power of the signal increases. In part b the consequences of poor frequency estimation are even more serious. For this experiment, S_2 is at a frequency of 56.25% of the Nyquist frequency and has a power of 0.1. Notice that the gain is 1 for moderate powers of S_1 and we see no effects from poor filter design. However, at high input power levels the gain drops to 0 because the poor estimate of the power spectrum erroneously gives the power at S_2's frequency as more than 1, even though it is really 0.1!

A pertinent question at this point is how contrived are these results? Actually the frequencies used in these experiments were carefully selected: 25% and 56.25% of the Nyquist frequency correspond to integer bin numbers of the 32 point DFT used in the experiment. 28.125% and 59.375% of the Nyquist frequency are frequencies exactly midway between integer bin numbers. The frequencies between integer bin numbers maximize the undesirable effects. In fact, if the length of the DFT used is doubled, all of the signals would correspond to integer bin numbers and the system would show no undesirable characteristics during these experiments. However,



PART a



PART b

Experiment 2 using the DFT/threshold based power separator

Figure 2.10

no matter how long the DFT is there will always be a range of frequencies which do not correspond to integer bin numbers. Any signal with a non-integer bin number will cause poor system performance. The exact gain levels and thresholds will, of course, depend on the actual frequencies, but for any length of DFT there will be many choices of signal frequency for which the power separator will perform poorly with results similar to those shown in Figures 2.9 and 2.10.

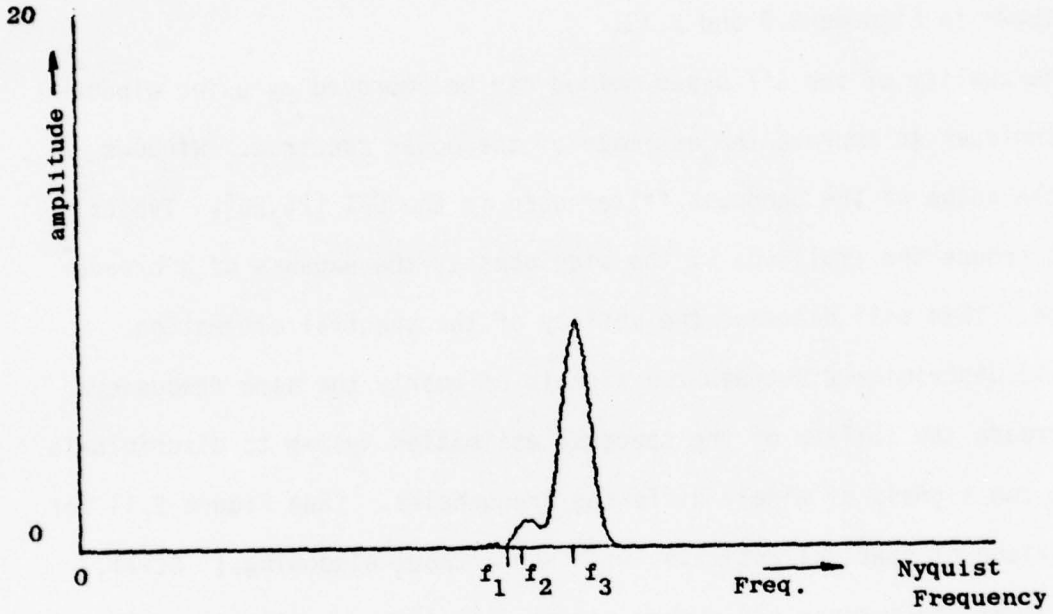
The quality of the DFT based method can be improved by using windowing techniques to improve the estimate of the power spectrum. Windows alter the shape of the bandpass filter used in the DFT [25,26]. Typical windows reduce the amplitude of the sidelobes at the expense of a broader mainlobe. This will decrease the ability of the spectral estimation system to discriminate between two signals of nearly the same frequency, and increase the ability of the spectral estimation system to discriminate between two signals of widely differing frequencies. (See Figure 2.11 for a comparison of spectral estimates with and without windowing.) Still, a high-power non-integer bin number signal will cause phantom power to appear throughout the spectral estimate. This 'spillover' or 'leakage' will result in poor performance unless the input signal powers are known to be reasonably bounded. If the signal powers are bounded to a small range, this technique could be useful, especially since fast, off-the-shelf spectral analysis hardware exists.

2.5 An Adaptive Power Separator with both Soft and Hard Constraints

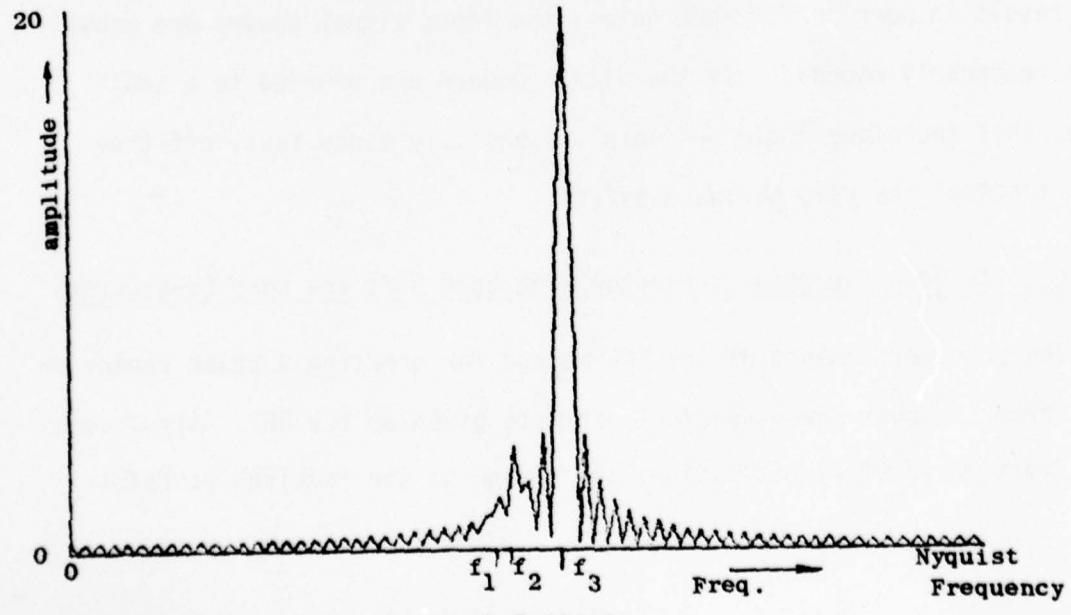
The poor performance of the DFT method for creating a power separator stems from the poor power spectral estimate given by the DFT. Also, even with improved spectral estimation, the design of the required bandstop

COMPARISON OF WINDOWED SPECTRAL ESTIMATION

Signal 1: Frequency = f_1; Amplitude = 1
Signal 2: Frequency = f_2; Amplitude = 3
Signal 3: Frequency = f_3; Amplitude = 20



Windowed Spectral Estimate
(using Blackman Window)



Raw Spectral Estimate

Figure 2.11 -56-

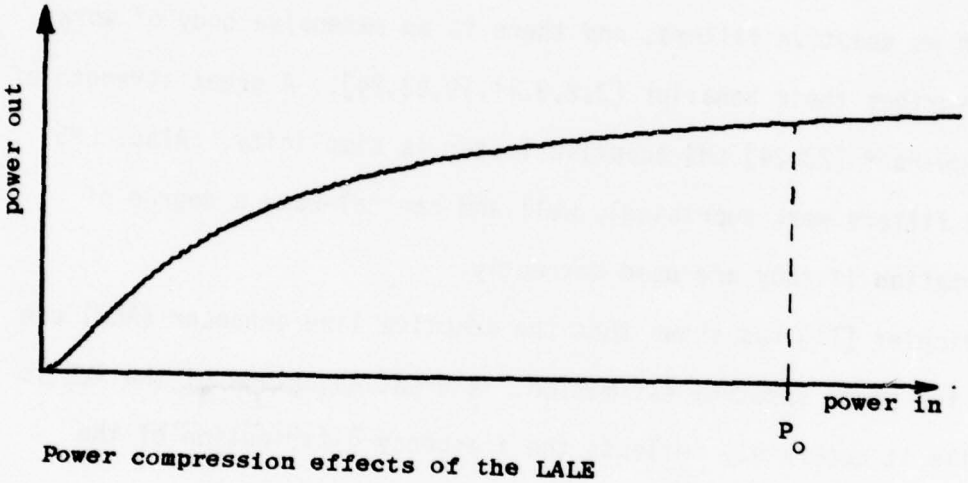
filters would be troublesome because if an FIR filter is used, only a finite number of zeros are implemented. For good signal rejection, at least one of the zeros should lie exactly on the signal frequency. This requires accurate frequency estimation and accurate filter design. Thus a static design, although theoretically feasible, may fail to work well due to minor errors such as roundoff during computation. An adaptive system, which can examine its performance and modify its parameters to improve its performance, will often avoid numerical difficulties and thus is clearly desirable.

Adaptive systems can be used for on-line filter design. Such filters are known as adaptive filters, and there is an extensive body of work which describes their behavior [3,8,9,11,19,23,24]. A great strength of the Widrow-Hoff [23,24] LMS adaptive filter is simplicity. Also, LMS adaptive filters work suprisingly well and can tolerate a degree of misinformation if they are used correctly.

Treichler [11] has shown that the adaptive line enhancer (ALE) can be used for power spectral estimation. A great advantage of the ALE is that while it accurately reflects the frequency distribution of the spectrum, it compresses the power axis (see Figure 2.12).

$$P_{out} = \frac{P_{in}}{\left(\frac{2}{n}(P_n+\gamma) + P_{in}\right)^2}$$

Because of this, all signals with power greater than P_o appear to have about the power of a signal with power P_o. The power-out versus power-in curve is monotone, however, and can thus be used to decide which signals are above a selected threshold. Therefore the ALE creates a one-to-one nonlinear distorted estimate of the power spectrum which is



Power compression effects of the LALE

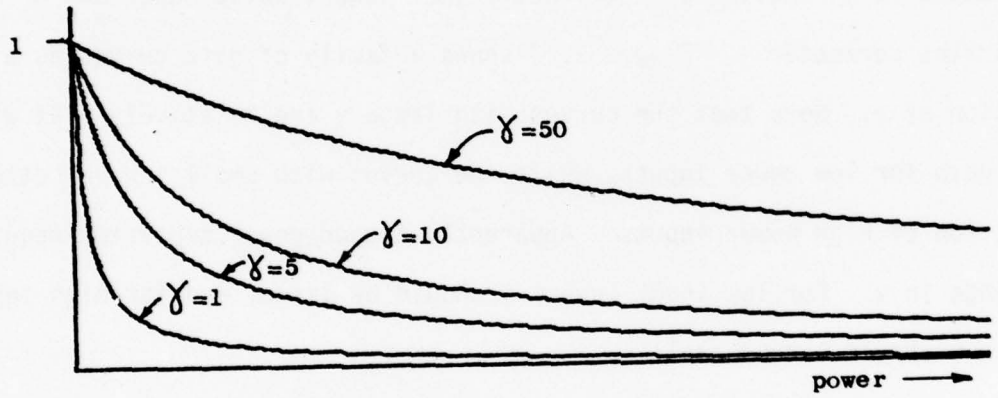
Figure 2.12

less susceptible to spillover, due to the power compression, but can still be used to decide which frequencies exceed a predetermined power level.

Two other advantages of using adaptive filters are: 1) if used properly, adaptive filters can partially accommodate incorrect information 2) the addition of soft constraints to the ALE provides a measure of power separation. By careful attention to the advantages of adaptive filters a relatively simple power separator can be devised.

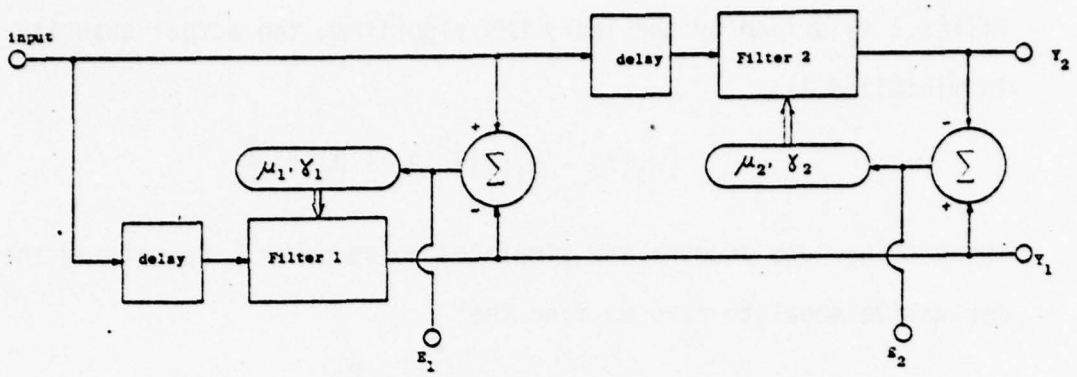
The design starts with a leaky (soft constrained) ALE (see Figure 2.3). The output is a function of the input signal power, noise power and a constraint parameter γ. Figure 2.13 shows a family of gain curves as a function of γ. Note that the curves with large γ are relatively flat with high gain for low power inputs, while the curves with small γ have better rejection of high power inputs. Apparently a good power separator requires a change in γ. For low input powers γ should be large, and for high input powers γ should become small.

Figure 2.14.B shows the gain curves from the input to the various outputs of the system shown in Figure 2.14.A. Filter 1 is an ordinary LALE with outputs Y_1 and E_1. Filter 2 is a LALE with a scaled desired response, that is, instead of the desired response to filter 2 being the input signal, the desired response to filter 2 is the input signal filtered by filter 1. Thus the desired response signal to filter 2, at any particular frequency f, is scaled by the gain of filter 1 at the frequency f. Since filter 2 is an adaptive filter which tries to minimize the average power of the error output, it will filter the input signal (power P_1) in such a way as to minimize the power of $[H_2(f) - H_1(f)]^2 P_1$ (where $H_2(f)$ is the gain of filter 2 at frequency f). However, since

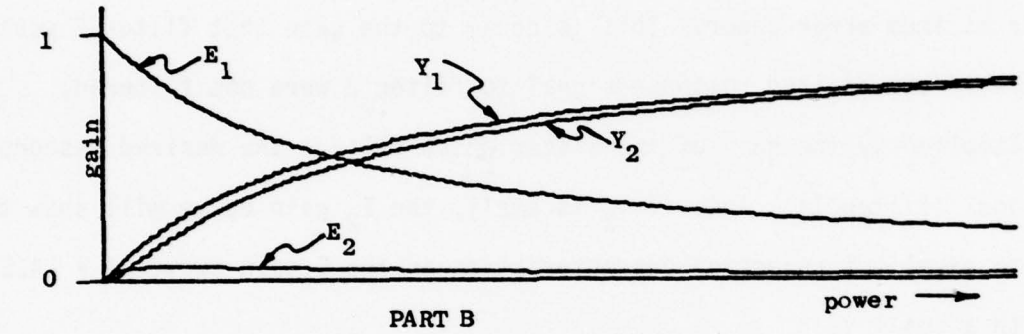


Gain of the LALE as the constraint γ is varied

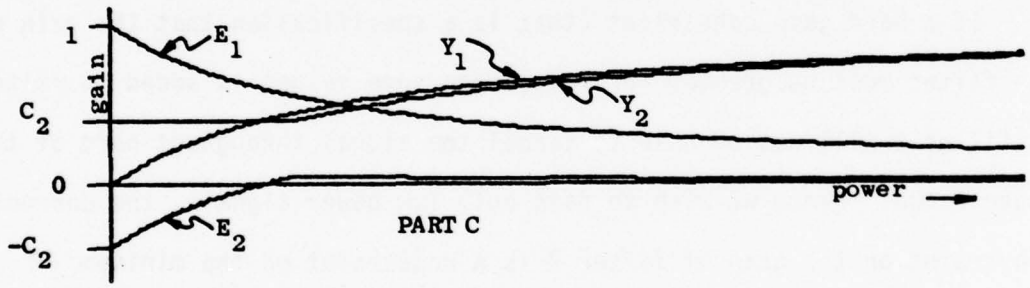
Figure 2.13



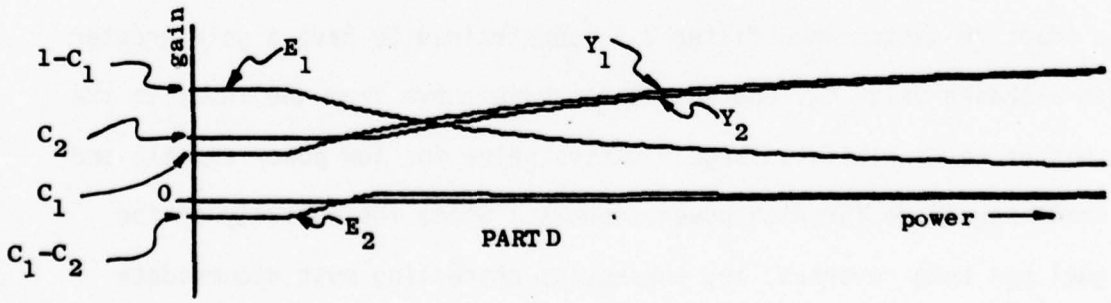
PART A



PART B



PART C



PART D

Development of improved power separator

Figure 2.14 -61-

filter 2 is driven by the leaky LMS algorithm, the actual quantity that is minimized is:

$$[H_2(f) - H_1(f)]^2 P_1 + H_2(f) \frac{2}{n} \gamma$$

which takes into account the simulated noise power. By setting the derivative equal to zero we find that

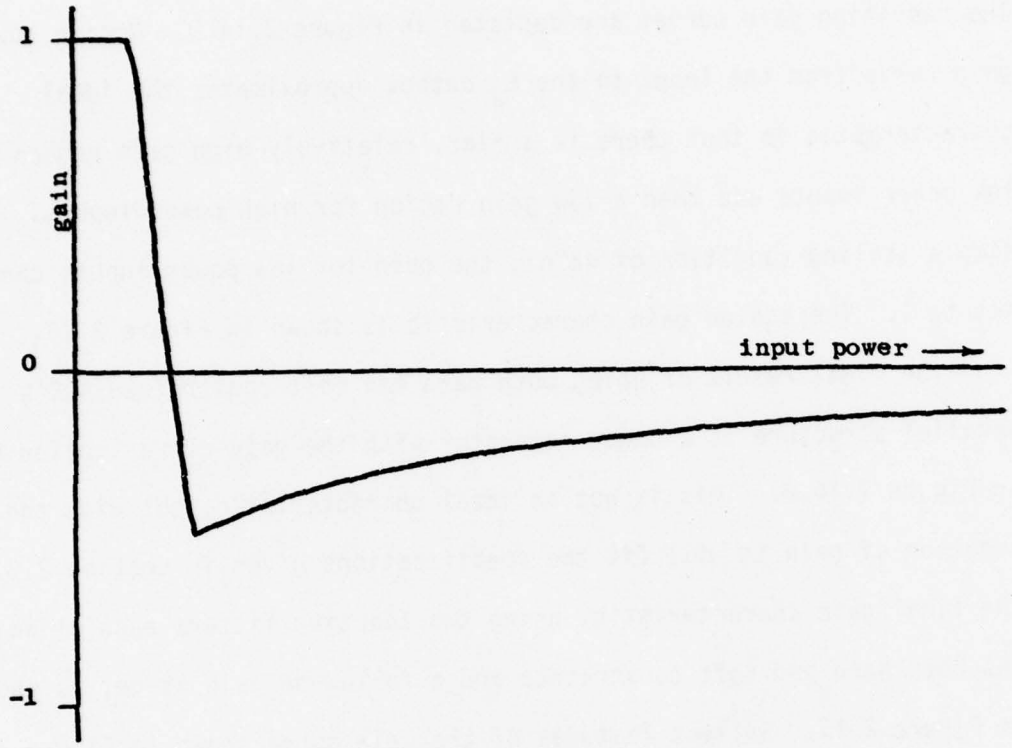
$$H_2(f) = H_1(f) \frac{P_1}{P_1 + \frac{2}{n} \gamma}$$

for minimum error power. This is equal to the gain that filter 2 would have if the desired response signal to filter 2 were not filtered, multiplied by the gain of the filter which filters the desired response signal (filter 1). Thus if γ_2 is small, the E_2 gain curve will show the same excellent rejection characteristics as the E gain curve of a LALE with a small γ.

If a hard gain constraint (that is a specification that the gain of the filter must be greater (or less) than some value) is added to filter 2, filter 2 will not be able to cancel the signal throughout part of the power range. Since we wish to pass only low power signals, the correct constraint on the gain of filter 2 is a constraint on the minimum allowed gain at any frequency. Figure 2.14.C shows the gain curves of the adaptive system when filter 2 is constrained to have a gain greater than a chosen value C_2. Note that the gain curve from the input to the E_2 output now exhibits a large negative value for low power signals and a small magnitude for high power signals. Since the polarity of the signal has been reversed, any subsequent processing must accommodate polarity changes.

The final imperfection, rapid decay of the gain in the lower power region, can be alleviated by constraining the minimum gain of filter 1. The resulting gain curves are depicted in Figure 2.14.D. Notice how the gain curve from the input to the E_2 output approximates the ideal characteristic in that there is a flat, relatively high gain region for low power inputs and then a low gain region for high power inputs. By simple scaling (addition of gain), the gain for low power inputs can be set to 1. The scaled gain characteristic is shown in Figure 2.15.

The final result of using both hard and soft constrained ALE's in a parallel structure is a power separator with the gain curve labeled E_2 in Figure 2.14.D. This is not an ideal characteristic, but with the addition of gain it will fit the specifications given in section 2.3. The final gain characteristic, using two adaptive filters each of which has both hard and soft constraints and a follow-on gain stage, is shown in Figure 2.15. Salient features of the gain curve shown in Figure 2.15 are: a gain of 1 for low power inputs, a transition region in which the gain decreases rapidly, and a low gain for high power inputs. Thus the filter depicted in Figure 2.14.A is a power separator. However, while the theory of soft-constrained adaptive filters is well developed, there is no corresponding theory for the hard constraints that were assumed in this discussion. Several methods suggest themselves. Each has drawbacks, each has advantages. The final quality of the power separator will depend on how well the hard constraints can be implemented. One method is to use a DFT of the LALE weight vector to estimate the gain of the filter. If the gain is too low, the weight vector is appropriately modified to increase the gain. The limitations of the DFT (finite resolution and spillover) will apply to this technique; however, the



Scaled gain curve from input to E_2 output

Figure 2.15

nonlinear gain characteristics of the LALE will mitigate the spillover effects of the DFT. The simplicity of this method, especially since the DFT of the LALE weight vector is already available in certain implementation techniques, may outweigh slight imperfections.

Alternatively, more exact methods of determining the minimum gain and then appropriately altering the weight vector may have to be used. Such methods as determining the minimal points by solving for the roots of the derivative of the gain polynomial, or, simply evaluating the gain on a fine net of points and searching for low areas might be used. These methods all require increased computational effort, so some judgement will have to be made which weighs the relative demerits of each technique.

2.6 A Simplified Analysis with Sinusoidal Inputs

This section will analyze the response of a single, DFT-constrained, LALE (see Figure 2.16) when the input consists of sinusoids at the DFT bin frequencies. Sinusoids are used because Treichler [11] showed that narrowband signals with a bandwidth less that 1/nT could be modeled as sinusoids when analyzing the ALE. Furthermore, by restricting attention to sinusoids we will be able to make general statements about the response of the LALE. That is, we will be able to derive equations which relate the gain and frequency characteristics of the LALE to the power and frequency of the individual inputs. If we analyze broadband inputs, we find that the solutions are complex and involve interactions between all of the inputs. Consequently, when the input consists of sinusoids, we can say what the response to a particular input will be without having to analyze the other inputs. On the other hand, when the input consists of broadband inputs, we have to know all of the inputs in detail before we can determine the response to a single input. The advantages of making

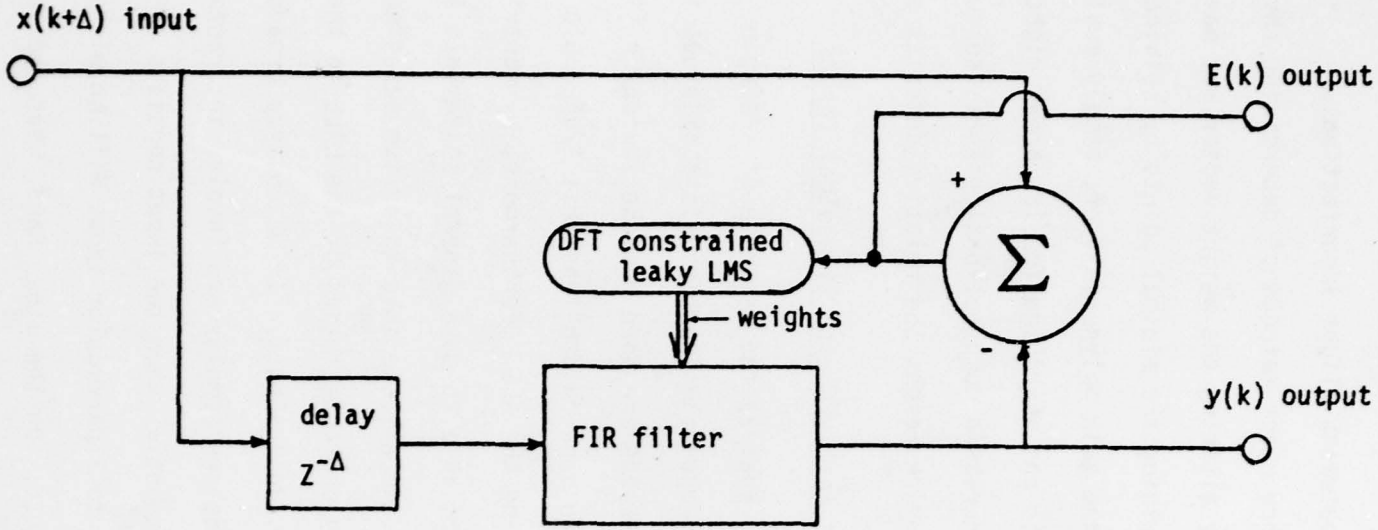

x(k+Δ) input
E(k) output
+
DFT constrained
leaky LMS
Σ
weights
-
delay
$Z^{-\Delta}$
FIR filter
y(k) output

Figure 2.16

general statements about a class of useful signals (narrowband), as contrasted to analyzing, case-by-case, other more general signals, cause this work to concentrate on sinusoidal inputs. This section further restricts the input by requiring the sinusoids to be at DFT "bin" frequencies (strictly speaking, we mean integer bin number when we say at or on a bin frequency, i.e. some frequency which is an integer multiple of 1/nT). Sinusoids which are at a DFT bin frequency are eigenvectors of the DFT. This greatly simplifies the analysis of the constraint, which in turn simplifies the development of the DFT constrained leaky adaptive line enhancer (DLALE).

The purpose of the hard constraint is to prevent the gain of the filter from being less than a selected threshold. The DFT constraint uses the DFT of the weight vector to assess the gain of the filter.

CONSEQUENCE: The DFT of the weight vector of an FIR filter is the gain of the FIR filter.

DEMONSTRATION: The transfer function of an FIR filter is:

$$H(z) = \sum_{i=0}^{n-1} [W]_i \, Z^{-i} \quad .$$

Where: W is the vector of weighting coefficients

$$Z = e^{sT} \quad .$$

To find the frequency response at a particular frequency:

$$f = \frac{m}{n} \Omega$$

where Ω = sampling frequency = 1/T

m = bin number

we substitute $e^{\frac{j2\pi m}{n}} = Z$ into H(z) which yields:

$$H(m) = \sum_{i=0}^{n-1} [W]_i \, e^{-j\frac{2\pi}{n} im} \qquad 0 \leq m \leq n-1 \quad . \qquad 2.5$$

This is the expression for the DFT of W, the weight vector.

Even though the DFT is a continuous function of frequency, in practice the DFT is evaluated only at n distinct frequencies. These frequencies, which are multiples of 1/nT, serve to represent the entire response since any other value can be determined by interpolation. Thus the DFT constraint uses a sampled version of the gain of the FIR filter. The ideal constraint requires that the gain of the filter be greater than some minimum value C. A plausible scheme for implementing the constraint, given the DFT of the weight vector (we will use the term gain vector to refer to the DFT of the weight vector), is to examine the gain vector and to set any value which is less than the threshold to the value of the threshold. Then a new weight vector can be derived as the inverse DFT of this modified gain vector. Obviously the new weight vector will meet the constraint criterion at the n DFT frequencies which correspond to integer bin numbers. The only question still unanswered is the phase to use when constraining the gain. Since the phase of the ideal prediction filter is known, it is reasonable to use the ideal phase advance of 2π m / n. (This has far-reaching implications when the sinusoid is not at an integer bin number.) Because the weight vector is real, the DFT of the weight vector will be Hermitian [27]. Similarly, the modified gain vector must be Hermitian to guarantee a real weight vector.

A DFT constrained LLMS adaptive filter (DLLMS), employing the criteria mentioned above, is depicted in Figure 2.17. If we define F(W) to be the DFT of the vector W, and $[W]_i$ to be the i^{th} element of the vector W, then the constraint can be expressed mathematically as:

$$G = \text{gain vector} = F(W) \qquad 2.6.A$$

$$[Gm]_i = [\text{modified gain vector}]_i = \begin{cases} [G]_i & \text{if } [G]_i > C \\ Ce^{j\,2\pi\,i/n} & \text{if } [G]_i < C \end{cases} \qquad 2.6.B$$

$$W_c = \text{Constrained weight vector} = F^{-1}(Gm) \ . \qquad 2.6.C$$

This defines the operation of the DFT based constraint.

We now define the adaptive algorithm by which a new weight vector is derived. This algorithm is known as the LLMS (leaky LMS) algorithm:

$$W(k+1) = \nu W(k) + 2\mu e(k)\, X(k) \qquad 2.7.A$$

$$e(k) = d(k) - y(k) = d(k) - W^t(k)\, X(k) \qquad 2.7.B$$

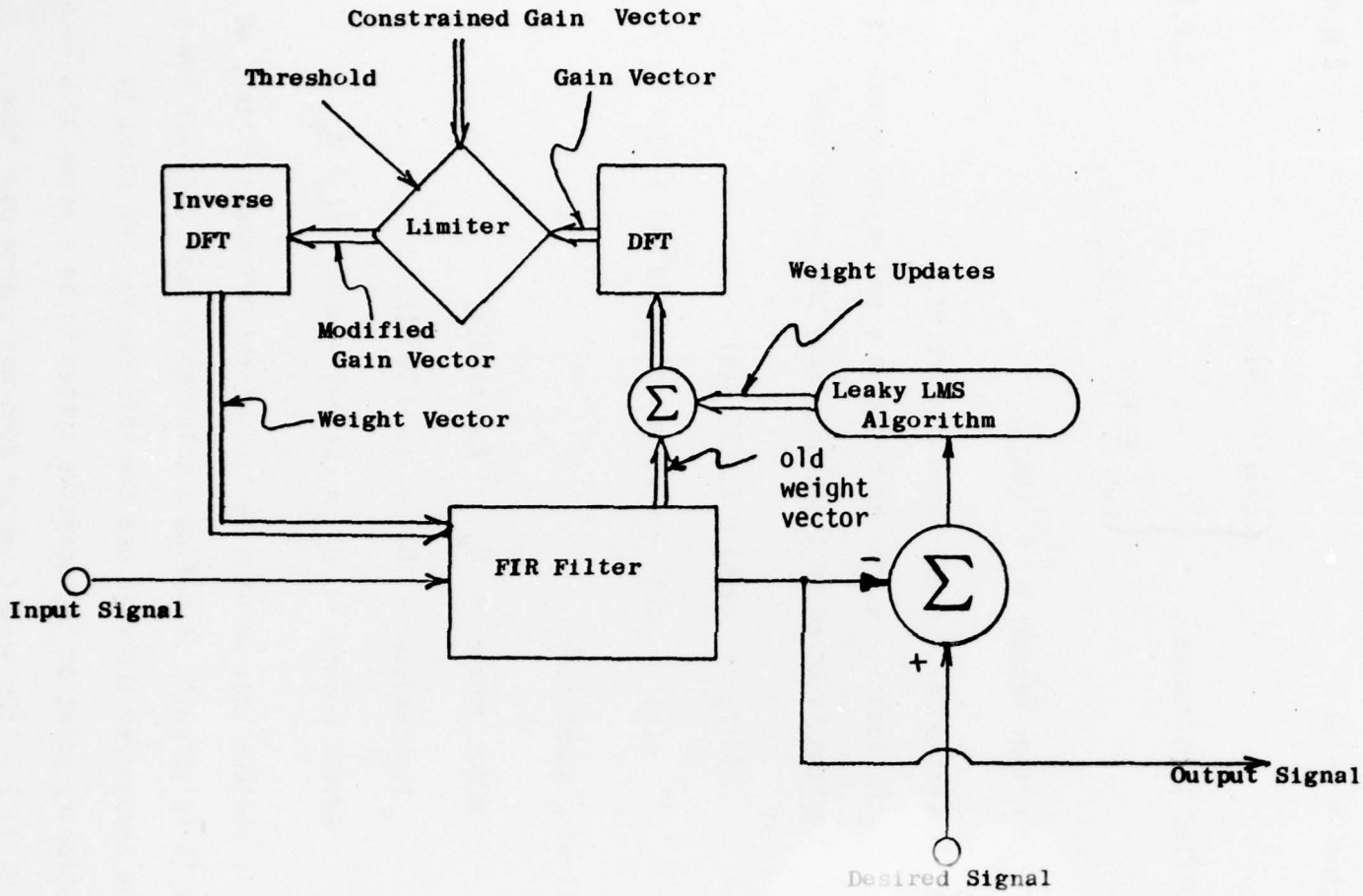
Define the following quantities:

$$\text{error update} = U_e = 2\mu e(k)\, X(k)$$

$$\text{leak update} = U_l = (\nu - 1)\, W(k)$$

$$\text{weight update} = U_w = W(k+1) - W(k) = U_l + U_e$$

While it is obvious that we wish to 'solve' the DFT based filter, we need to define 'a solution.' To define a solution we must consider how a solution can be recognized and what use the solution is. We start by defining a stationary point of the adaptive system to be a value of W such that W(k+1) = W(k), i.e. the weight vector does not change over time.



DFT Constrained Leaky LMS Filter

Figure 2.17

Also, a stable stationary point is a stationary point at which any perturbation, dW, from the stationary point results in changes which cause W to return to the stationary point. Stable stationary points are interesting solutions to the DLLMS weight vector because once the DLLMS weight vector gets near to a stable stationary point it will progress to the stationary point and stay there indefinitely. Thus we will endeavor to find stationary points of the DLLMS filter and then check for stability and uniqueness.

Given that $x(k) = a\,\mathrm{Cos}(\omega_o kT + \phi)$, where ω_o corresponds to a bin number of m_o, we wish to find the stationary points of the DLLMS filter. Changing to phasor notation with an implied frequency of ω_o, the input is:

$$x = a\angle\phi$$

therefore

$$y = a\angle\phi\,[F\{W\}]_{m_o} \tag{2.8}$$

$$\begin{aligned} e &= a\angle\phi\ \sqrt{\omega_o}^{T} - a\angle\phi\,[F\ W\ ]_{m_o} \\ &= a\angle\phi\ (1\angle\sqrt{\omega_o}^{T} - [F\{W\}]_{m_o}) \\ &= a\angle\phi\ \ g\angle\theta \end{aligned} \tag{2.9}$$

from this we can compute:

$$\begin{aligned} [U_e]_i &= 2\mu e(k)[X(k)]_i \\ &= 2\mu a\ g\ \mathrm{Cos}(\omega_o T+\theta+\phi)\ a\ \mathrm{Cos}(\omega_o(k-i)T+\phi) \\ &= 2\mu a^2 g\ \tfrac{1}{2}\,[\mathrm{Cos}(\theta+\omega_o iT) + \mathrm{Cos}(2\omega_o kT+\theta-\omega_o iT+2\phi)] \\ &= \mu a^2 g[\mathrm{Cos}(\theta+\omega_o iT) + \mathrm{Cos}(2\omega_o kT+\theta-\omega_o iT+2\phi)] \quad . \end{aligned} \tag{2.10}$$

The first term of U_e is not a function of time (index k). The second term is at twice the input frequency. The LLMS weight update equation (Eq. 2.7.A) can be Z-transformed to give:

$$\frac{W(z)}{U_e(z)} = \frac{1}{z - \nu} \quad . \qquad 2.11$$

Of course $U_e(z)$ is a function of W(z). However, if the value of μ is small then W will be approximately constant which means that e, and hence U_e, will be approximately constant.

ASSUMPTION 1: μ is small

If the value of μ is small, then $W(k+\Delta) = W(k)$ + "leaky" integral of U_e. The integral of the first term of U_e is $1/(1-\nu)\ \mu a^2\ g\ \cos[\theta+\omega_o iT]$, which is large for $\nu \simeq 1$ (normal operation). The integral of the second term of U_e is a sinusoid at twice the input frequency (i.e. $\mathrm{Cos}[2\omega_o kT + \tau]$) multiplied by some gain factor.

ASSUMPTION 2: $0.1\,\Omega <= \omega_o <= 0.4\,\Omega$

If ASSUMPTION 2 is correct, then the gain for the second term of U_e will be very much less than the gain for the first term of U_e. Thus, as a simplifying approximation, the double frequency component of U_e is ignored in the following analysis.

Define:

$$U_{em} = \text{modified } U_e = \text{constant terms of } U_e$$

so

$$[U_{em}]_i = \mu a^2\ g\ \mathrm{Cos}[\omega_o iT + \theta] \qquad 2.12$$

Thus, to solve for stationary points:

$$
\begin{aligned}
0 &= U_w \tilde{=} U_1 + U_{em} \\
&= (\nu-1)W + U_{em} \\
&= F\{(\nu-1)W\} + F\{U_{em}\} \\
&= (\nu-1)\, F\{W\} + \mu a^2 g \frac{n}{2} \left(D(m_o)\, e^{j\theta} + D(-m_o)\, e^{-j\theta}\right)
\end{aligned}
\qquad 2.13
$$

$$
[D(m_o)]_i = \begin{cases} 0 & \text{if } i \neq m_o \\ 1 & \text{if } i = m_o \end{cases}
$$

Therefore

$$
[F\{W\}]_i = 0 \quad \text{if } i \neq m_o, -m_o \qquad 2.14
$$

$$
(1-\nu)[F\{W\}]_{m_o} = \mu a^2 \frac{n}{2} g\, e^{j\theta}
$$

$$
= \mu a^2 \frac{n}{2} \left(e^{j\omega_o T} - [F\{W\}]_{m_o}\right)
$$

$$
\Rightarrow [F\{W\}]_{m_o} \left(1 + \frac{\frac{n}{2}\frac{a^2}{2}}{\frac{1-\nu}{2\mu}}\right) = \frac{\frac{n}{2}\frac{a^2}{2}}{\frac{1-\nu}{2\mu}} e^{j\omega_o T}
$$

$$
\Rightarrow [F\{W\}]_{m_o} = \frac{\frac{n}{2} SNR'}{\frac{n}{2} SNR' + 1} e^{j\omega_o T} \qquad 2.15
$$

also

$$
[F\{W\}]_{-m_o} = \overline{[F\{W\}]_{m_o}} \quad . \qquad 2.16
$$

This expression, $[F(W)]_{m_o}$, for the gain of the FIR filter agrees with the expression Treichler [11] developed by a different approach.

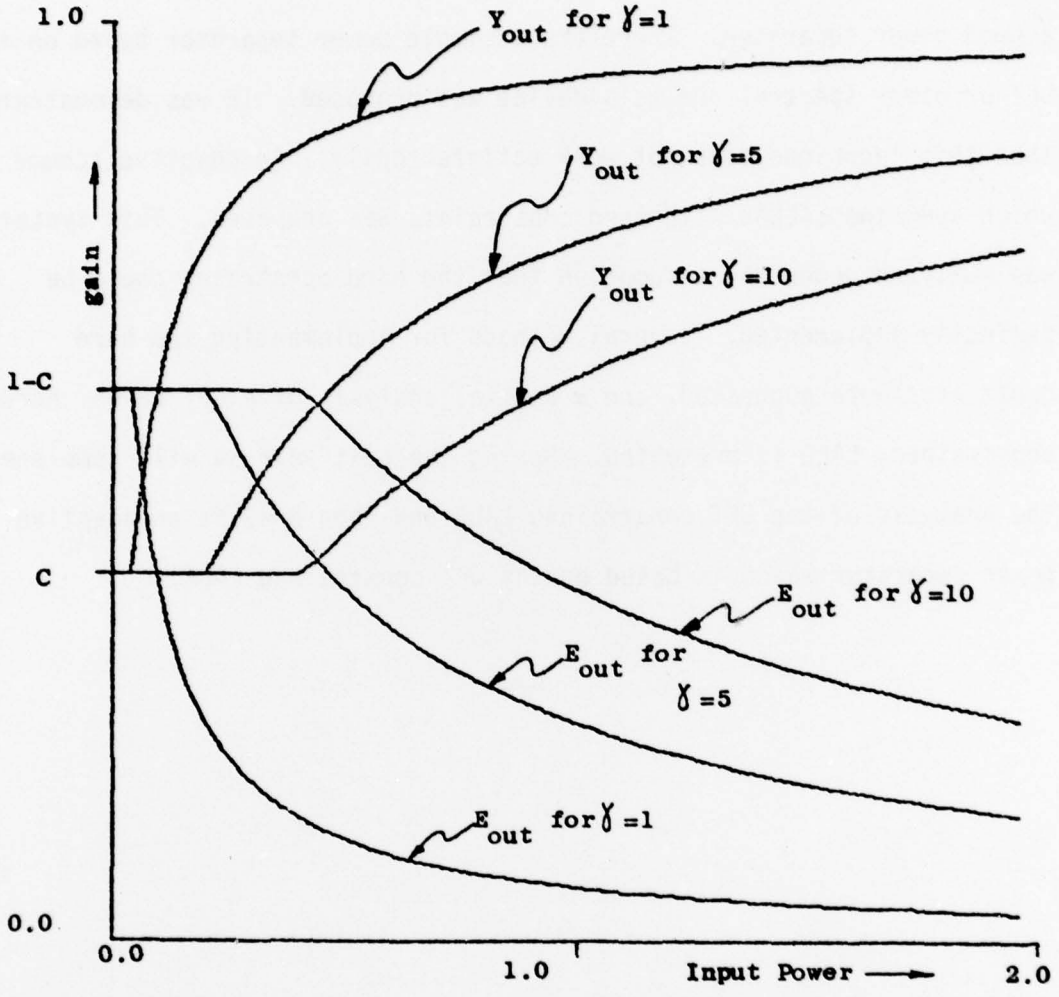
The expressions for F(W) (2.14, 2.15, 2.16) are not complete since they allow the gains to be less than the threshold value. The constraint will cause any $[F(W)]_i$ with a magnitude which is less than the threshold C to be set to the constrained value $[B]_i = C\, e^{j\frac{2\pi}{n}i}$. Applying the constraint to a particular bin does not affect the solution for any other bin because DFT bins correspond to eigenvectors of the FIR filter. Thus for the case of a sinusoid at a bin frequency of m_o, the gain of the DLLMS filter at frequency m_o is:

$$g = \begin{cases} Ce^{j\omega_o T} & \text{if } \dfrac{\frac{n}{2}SNR'}{1+\frac{n}{2}SNR'} < C \\ \dfrac{\frac{n}{2}SNR'}{1+\frac{n}{2}SNR'}\, e^{j\omega_o T} & \text{if } \dfrac{\frac{n}{2}SNR'}{1+\frac{n}{2}SNR'} \geq C \end{cases} .$$

Furthermore, the DLALE will respond independently to different sinusoids provided that they are all at bin frequencies. DLALE gain curves for this case are shown in Figure 2.18. (The gain curves of Figure 2.18 are labelled in correspondence with Figure 2.16.)

2.7 Conclusions

In this section of the final report we have discussed the concept of power separation. A power separator is a system which passes, unchanged, low power signals and severely attenuates high power signals. Such a system can obviously be used to protect a system against high



DLALE gain curves for on-bin sinusoids
(Gain is measured from input to designated output)

Figure 2.18

power jamming signals, but we have also outlined other possible uses. We have also presented a practical gain specification for a power separator, and compared the leaky adaptive line enhancer (LALE) to the specified gain characteristic. Since the LALE has a slow rolloff it does not make a good power separator. Therefore a simple power separator based on a DFT or other spectral analysis device was proposed. It was demonstrated that this technique does not work satisfactorily. An adaptive scheme which uses two LALE's with hard constraints was proposed. This system was analyzed under the assumption that the hard constraint could be perfectly implemented. Several methods for implementing the hard constraint were suggested, and a partial analysis of a DFT based, hard constrained, LALE is presented. During the next year we will complete the analysis of the DFT constrained LALE and then analyze an adaptive power separator which is based on the DFT constrained LALE.

BIBLIOGRAPHY

1. B. Widrow, "Adaptive Filters," in Aspects of Network and System Theory, R. E. Kalman and N. DeClaris, Eds. New York: Holt, Reinhart, and Winston, Inc. 1970, pp. 563-587.

2. B. Widrow and J. M. McCool, "A Comparison of Adaptive Algorithms Based on the Methods of Steepest Descent and Random Search," IEEE Trans. on Antennas and Propagation, Vol. AP-24, September 1976, pp. 615-637.

3. B. Widrow, et al, "Adaptive Noise Cancelling: Principles and Applications," Proceedings of the IEEE, Vol. 63, No. 12, December 1975, pp. 1692-1716.

4. J. T. Treichler, "Transient and Convergent Behavior of the Adaptive Line Enhancer," IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-27, No. 1, February 1979, pp. 53-62.

5. D. R. Morgan and S. E. Craig, "Real Time Adaptive Linear Prediction using the Least Mean Square Gradient Algorithm," IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-24, No. 6, December 1976, pp. 494-507.

6. R. L. Riegler and R. T. Compton, Jr., "An Adaptive Array for Interference Rejection," Proceedings of the IEEE, Vol. 61, No. 6, June 1973, pp. 748-758.

7. R. D. Gitlin, J. E. Mazo, and M. G. Taylor, "On the Design of Gradient Algorithms for Digitally Implemented Adaptive Filters," IEEE Trans. on Circuit Theory, Vol. CT-20, No. 2, March 1973, pp. 125-136.

8. L. J. Griffiths, "A Simple Adaptive Algorithm for Real-Time Processing in Antenna Arrays," Proceedings of the IEEE, Vol. 57, No. 10, October 1969, pp. 1696-1704.

9. O. L. Frost III, "An Algorithm for Linearly Constrained Adaptive Array Processing," Proceedings of the IEEE, Vol. 60, No. 8, August 1972, pp. 926-935.

10. C. S. Williams, "Adaptive Channel Modeling for Digital Communication in the Presence of Multipath and Doppler," Ph.D. Dissertation, Department of Electrical Engineering, Stanford University, 1977.

11. J. T. Treichler, "The Spectral Line Enhancer - The Concept, an Implementation, and an Application," Ph.D. Dissertation, Dept. of Electrical Engineering, Stanford University, 1977.

12. N. Ahmed, G. R. Elliot, and S. D. Stearns, "Long Term Instability Problems in Adaptive Noise Cancellers," Sandia Report SAND78-1032, Sandia Laboratories, Albuquerque, NM, 87185, August 1978.

13. W. D. White, "Artificial Noise in Adaptive Arrays," IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-14, No. 2, March 1978, pp. 380-384.

14. C. L. Zahm, "Application of Adaptive Arrays to Suppress Strong Jammers in the Presence of Weak Signals," IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-9, No. 2, March 1973, pp. 260-271.

15. H. L. Van Trees, Detection, Estimation, and Modulation Theory, Part I, New York: John Wiley and Sons, 1968.

16. A. P. Sage and J. L. Melsa, Estimation Theory with Applications to Communication and Control, New York: McGraw-Hill, 1971.

17. T. Kailath, Lectures on Linear Least Squares Estimation, New York: Springer Verlag, 1976.

18. D. J. Wilde, Optimum Seeking Methods, Englewood Cliffs, NJ: Prentice-Hall, 1964.

19. S. P. Applebaum, "Adaptive Arrays," Special Projects Lab., Syracuse University Res. Corp., Rep. SPL 769.

20. R. S. Simpson and R. C. Houts, Fundamentals of Analog and Digital Communication Systems, Allyn and Bacon, Inc., Boston, 1971.

21. B. Widrow, R. Chestek and J. R. Treichler, "Research on Adaptive Antenna Techniques: Final Report," Stanford University, Stanford, March 1977, Naval Air Systems Command Contract N00019-76-C-0250.

22. B. Widrow, R. Chestek, T. Saxe, "Research on Adaptive Antenna Techniques II: Final Report," Stanford University, Stanford, March 1978, Naval Air Systems Command Contract N00019-77-C-0194.

23. B. Widrow and M. Hoff, Jr., "Adaptive Switching Circuits," in IRE WESCON Conv. Rec., Pt. 4, pp. 96-104, 1960.

24. B. Widrow, "Adaptive Filters I: Fundamentals," Stanford Electronics Labs. Technical Report No. 6764-6, December 1966.

25. A. Papoulis, Signal Analysis, McGraw-Hill Book Co., New York, 1977.

26. A. V. Oppenheim and R. W. Schafer, Digital Signal Processing, Prentice-Hall, Inc., New Jersey, 1976.

27. R. Bracewell, The Fourier Transform and Its Applications, McGraw-Hill Book Co., New York, 1965.